# Kapitel 1

# Mengen und Funktionen

## 1.1 Wir nehmen den naiven Standpunkt in der Mengenlehre ein ...

... und wollen uns nicht um die allgemeinst-mögliche Konstruktion von Mengen kümmern. [1] Wir bemerken, daß die Definition einer Menge, die von Georg Cantor gegeben wurde, nämlich

> Unter einer Menge verstehen wir jede Zusammenfassung $M$ von bestimmten wohlunterschiedenen Objekten in unserer Anschauung oder unseres Denkens (welche die Elemente von $M$ genannt werden) zu einem Ganzen.

und die wohl heutzutage in der einen oder anderen Form im Hinterkopf eines jeden Studenten herumspukt, zu allgemein ist, um keine Widersprüche aufkommen zu lassen. Cantors Definition betont aber zwei wichtige Aspekte:

- **bestimmt:** Zu einem „Objekt“ $a$ und einer Menge $M$ kann genau eine der zwei Möglichkeiten:
  - entweder $a \in M$, dh. $a$ ist Element von $M$
  - oder $a \notin M$, dh. $a$ ist kein Element von $M$

  eintreten – es gibt also kein „vielleicht“ oder „ich weiß es nicht“. Damit wird $a \in M$ (und genauso $a \notin M$) zu einer **Aussage**, dh. zu einem Ausdruck, dem man genau einen Wahrheitsgehalt - also entweder „WAHR“ oder „FALSCH“ zuordnen kann.

- **wohlunterschieden:** Eine Menge $M$ liefert eine Vorschrift $=$, wann Elemente $a, b \in M$ (in $M$) gleich oder verschieden sind; es gilt also für $a, b \in M$ wieder genau eine der Beziehungen:
  - entweder $a = b$, dh. $a$ und $b$ bezeichnen in $M$ das gleiche Element
  - oder $a \neq b$, dh. $a$ und $b$ bezeichnen verschiedene Elemente von $M$.

[1] Wir werden in den nächsten Kapiteln die Axiome zur Konstruktion der natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, ...\}$ angeben und dann daraus die ganzen Zahlen $\mathbb{Z}$, die rationalen Zahlen $\mathbb{Q}$, die reellen Zahlen $\mathbb{R}$ und die komplexen Zahlen $\mathbb{C}$ konstruieren...

Ausgehend von Cantors Definition sind also zwei Mengen gleich, wenn sie dieselben Elemente enthalten – in der axiomatischen Mengenlehre ist das ein **Axiom**.

**Beispiel 1.1.1.**

- Eine Menge mit endlich vielen Elementen kann durch explizites Aufschreiben aller ihrer Elemente notiert werden, etwa $M = \{\diamondsuit, \heartsuit\}$. Die Aufzählung beginnt mit $\{$ und endet mit $\}$ – den **Mengenklammern**. Die Mengen $\{\diamondsuit, \square\}$ und $\{\square, \diamondsuit, \square\}$ enthalten dieselben Elemente (bei einer Aufzählung von Elementen einer Menge kommt es nicht auf die Reihenfolge oder Wiederholungen an!) und sind daher gleich.

- Vorsicht: Man sollte hier genau auf die Notation achten, sonst wird man verwirrt und meint für gleiche Objekte gibt es unterschiedliche Definitionen von „GLEICH“ und „VERSCHIEDEN“, aber das beschreibt dann verschiedene Situationen und führt zu verschiedenen Mengen. Ein Beispiel: Würfeln mit zwei Würfeln und zwar einmal mit einem roten und einem weißen Würfel und einmal mit zwei gleichen Würfeln. Würfelt man mit dem roten und weißen Würfel, so kann man etwa mit $(1, 2)$ notieren: Der rote Würfel zeigt eine Eins, der weiße Würfel zeigt eine Zwei. Die Menge der möglichen Ergebnisse wäre dann

$$\begin{aligned} M \;=\; \{&(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\\ &(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),\\ &(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\} \end{aligned}$$

  und zwei Elemente $(x_1, x_2), (y_1, y_2) \in M$ sind also gleich, wenn $x_1 = y_1$ und $x_2 = y_2$ sind (also der rote und der weiße Würfel dieselbe Zahl zeigen). Würfelt man mit zwei gleichen Würfeln, so sollte man eine andere Notation wählen, etwa $\langle 1; 2\rangle$ um das Ereignis „Es wurde eine Eins und eine Zwei gewürfelt“ zu notieren. Die Menge der möglichen Ergebnisse wäre hier

$$\begin{aligned} N \;=\; \{&\langle 1;1\rangle,\langle 1;2\rangle,\langle 1;3\rangle,\langle 1;4\rangle,\langle 1;5\rangle,\langle 1;6\rangle,\langle 2;2\rangle,\langle 2;3\rangle,\langle 2;4\rangle,\langle 2;5\rangle,\langle 2;6\rangle,\\ &\langle 3;3\rangle,\langle 3;4\rangle,\langle 3;5\rangle,\langle 3;6\rangle,\langle 4;4\rangle,\langle 4;5\rangle,\langle 4;6\rangle,\langle 5;5\rangle,\langle 5;6\rangle,\langle 6;6\rangle\} \end{aligned}$$

  wobei hier etwa $\langle 1; 2\rangle := \{(1, 2), (2, 1)\}$ die Menge mit den Elementen $(1, 2)$ und $(2, 1)$ ist und daher schon Identitäten wie $\langle 1; 2\rangle = \langle 2; 1\rangle$ in der Aufzählung berücksichtigt sind.

- Ist bereits eine Menge gegeben, dann lassen sich durch Angabe von bestimmten Eigenschaften Teilmengen definieren, zB. durch

$$\{n \in \mathbb{N} : n \text{ ist gerade }\}.$$

  Hier soll die Menge $\mathbb{N}$ der natürlichen Zahlen bereits bekannt sein und „$n$ ist gerade“ ist eine Eigenschaft, die die Teilmenge der geraden Zahlen definiert.

**Bemerkung 1.1.2.** Das Russelsche Paradoxon betrachtet die „Menge $R$ aller Mengen, die sich nicht selbst als Element enthalten“ und stellt die Frage, ob sich $R$ als Element enthält. Dies führt zu einem Widerspruch, denn wenn die Antwort „NEIN“ lautet, dann ist $R \in R$ – nach der Definition von $R$; lautet die Antwort aber „JA“ , dann gilt $R \notin R$ – wieder nach der Definition von $R$. Der Ausweg aus diesem Paradoxon liegt darin, daß es keine Menge gibt, die durch diese Eigenschaft gegeben wird. Die axiomatische Mengenlehre beschäftigt sich damit, Axiome anzugeben, mit deren Hilfe eine widerspruchsfreie Definition von Mengen möglich ist. Wir wollen jetzt dieses Programm nicht durchführen, sondern starten mit der Voraussetzung

Sei $X$ eine Menge...

und betrachten mögliche Konstruktionen mit dieser Menge $X$.

**Definition 1.1.3.** *Es seien $X$ und $Y$ Mengen.*

1. *$Y$ heißt eine* **Teilmenge** *von $X$, in Zeichen $Y \subseteq X$, wenn für alle $y \in Y$ auch $y \in X$ gilt.*
2. *$\mathcal{P}(X) := \{Y : Y \subseteq X\}$ heißt die* **Potenzmenge** *von $X$.*
3. *Ist $Y \subseteq X$, so heißt*

$$X \backslash Y := \{x \in X : x \notin Y\}$$

*das* **Komplement** *von $Y$ (in $X$).*

4. *$\emptyset$ bezeichnet die* **leere Menge***, dh. die Menge, die kein Element besitzt.*

**Beispiel 1.1.4.** Zum Beweis von $X = Y$ zeigt man oft $X \subseteq Y$ und $Y \subseteq X$; das braucht man zB. wenn es für eine Menge verschiedene Schreibweisen gibt, ein Beispiel:

$$X := \{2, 3\} = \{n \in \{2, 3, 4, 5\} : n^2 < 10\} =: Y.$$

In der ersten Form werden alle Elemente explizit aufgezählt (das ist nur bei Mengen mit endlich vielen Elementen möglich) – in der zweiten Form wird die Menge $Y$ (als Teilmenge von $\{2, 3, 4, 5\}$) durch die Angabe einer Eigenschaft (hier $n^2 < 10$) definiert. Der Nachweis von $X \subseteq Y$ erfolgt durch Berechnen von $2^2 = 4$ und $3^2 = 9$ was jeweils kleiner als 10 ist. Der Nachweis von $Y \subseteq X$ erfolgt durch Berechnen von $4^2 = 16 > 10$ und $5^2 = 25 > 10$, was $4 \notin Y$ und $5 \notin Y$ und damit $Y \subseteq X$ zeigt.

**Beispiel 1.1.5.** Für jede Menge $X$ gilt $\emptyset, X \in \mathcal{P}(X)$, etwa für $X = \{\diamondsuit, \heartsuit\}$ ist $\mathcal{P}(X) = \{\emptyset, \{\diamondsuit\}, \{\heartsuit\}, X\}$ und etwa $X \backslash \{\diamondsuit\} = \{\heartsuit\}$.

**Definition 1.1.6.** *Es seien $X$ und $I$ Mengen und es seien für alle $i \in I$ Teilmengen $X_i \subseteq X$ gegeben. Dann heißt*

$$\bigcup_{i \in I} X_i := \{x \in X : \text{ Es existiert ein } i \in I \text{ mit } x \in X_i \}$$

*die* **Vereinigung** *der Mengen $X_i, i \in I$ und*

$$\bigcap_{i \in I} X_i := \{x \in X : \text{ Für alle } i \in I \text{ gilt } x \in X_i\}$$

*der* **Durchschnitt** *der Mengen $X_i, i \in I$.*

**Beispiel# 1.** Ist $X = \{\diamond, \bullet, \diamondsuit, \heartsuit, \maltese, *, \odot, \triangle, \bigtriangledown, \infty\}$ und sind die Mengen $X_1 := \{\maltese, \odot, \infty\}$, $X_2 := \{\triangle, \infty, \heartsuit, *\}$ und $X_3 := \{\bullet, \diamond, *, \infty\}$ gegeben, dann ist

$$\begin{aligned} X_1 \cup X_2 &= \{\maltese, \odot, \infty, \triangle, \heartsuit, *\} \\ X_1 \cap X_2 &= \{\infty\} \\ X_1 \cup X_2 \cup X_3 &= \{\maltese, \odot, \infty, \triangle, \heartsuit, *, \bullet, \diamond, \} \end{aligned}$$

**Lemma 1.1.7.** *Es seien $X$ und $I$ Mengen, $X_i \subseteq X$ für $i \in I$, dann gilt:*

$$X \setminus \left( \bigcap_{i \in I} X_i \right) = \bigcup_{i \in I} (X \setminus X_i) \tag{1.1.1}$$

$$X \setminus \left( \bigcup_{i \in I} X_i \right) = \bigcap_{i \in I} (X \setminus X_i) \tag{1.1.2}$$

*Beweis.* Zum Beweis von (1.1.1):

„$\subseteq$“: Ist $x \in X \setminus \left( \bigcap_{i \in I} X_i \right)$, so gibt es (mindestens) ein $j \in I$ mit $x \notin X_j$, also $x \in X \setminus X_j$ und daher $x \in \bigcup_{i \in I} (X \setminus X_i)$, dh. $X \setminus \left( \bigcap_{i \in I} X_i \right) \subseteq \bigcup_{i \in I} (X \setminus X_i)$.

„$\supseteq$“: Für $x \in \bigcup_{i \in I} (X \setminus X_i)$ gibt es $j \in I$ mit $x \in X \setminus X_j$, also ist $x \notin \bigcap_{i \in I} X_i$, dh. $\bigcup_{i \in I} (X \setminus X_i) \subseteq X \setminus \left( \bigcap_{i \in I} X_i \right)$. $\square$

**Definition 1.1.8.** *Es sei $I$ eine Menge und $X_i$ sei für $i \in I$ eine Menge, dann heißt*

$$\prod_{i\in I} X_i := \{(x_i)_{i\in I} : x_i \in X_i \text{ für alle } i \in I\} \tag{1.1.3}$$

*das cartesische Produkt der Mengen $X_i, i \in I$. Die Gleichheit in $\prod_{i\in I} X_i$ ist folgendermaßen definiert: $(x_i)_{i\in I} = (y_i)_{i\in I}$ gilt in $\prod_{i\in I} X_i$ genau dann wenn $x_i = y_i$ für alle $i \in I$ erfüllt sind. Ist $X_i = X$ für alle $i \in I$, so schreibe $X^I$ statt $\prod_{i\in I} X_i$. Im Fall von endlich vielen Mengen $X_1, ..., X_n$ schreiben wir auch $X_1 \times ... \times X_n$ für das cartesische Produkt, bzw $X^n$ im Fall von $X = X_1 = ... = X_n$*

**Beispiel 1.1.9.**

- Schauen wir uns das einfachste Beispiel für ein cartesisches Produkt an; hier ist $I = \{1, 2\}$ eine Menge mit zwei Elementen. Für zwei Mengen $X_1$ und $X_2$ besteht das cartesische Produkt

  $$X_1 \times X_2 = \{(x_1, x_2) : x_1 \in X_1, x_2 \in X_2\}$$

  aus geordneten Tupeln mit zwei Einträgen. Zwei Elemente $(x_1, x_2) \in X_1 \times X_2$ und $(y_1, y_2) \in X_1 \times X_2$ sind genau dann gleich, wenn $x_1 = y_1$ (in $X_1$) und $x_2 = y_2$ (in $X_2$) gilt. Im Beispiel 1.1.1 haben wir beim Würfeln mit dem roten und dem weißen Würfel die Menge $X_1 = \{1, 2, 3, 4, 5, 6\}$ aller möglichen Ergebnisse des roten Würfels und die Menge $X_2 = \{1, 2, 3, 4, 5, 6\}$ aller möglichen Ergebnisse des weißen Würfels. Das cartesische Produkt $X_1 \times X_2 = \{1, 2, 3, 4, 5, 6\}^2$ gibt in diesem Fall die geordneten Ergebnisse (der erste Eintrag ist das Ergebnis des roten Würfels!) beim Würfeln mit zwei Würfeln an.

- Die Verallgemeinerung auf eine endliche Indexmenge $I$ geht bei konkreten Beispielen wie etwa bei $I = \{1, 2, 3\}$ und $X_1 = X_2 = X_3 = \mathbb{R}$ und

  $$\mathbb{R}^I = \{(x_1, x_2, x_3) : x_1 \in \mathbb{R}, x_2 \in \mathbb{R}, x_3 \in \mathbb{R}\}$$

  dann auch recht gut - beachte die Standardkonvention in diesem letzten Beispiel schreibt man $\mathbb{R}^3$ statt $\mathbb{R}^I$. Als Abkürzung kann man im letzten Beispiel auch $(x_i)_{i\in I}$ für ein Element $(x_1, x_2, x_3) \in \mathbb{R}^3$ schreiben. Die Schreibweise $(x_i)_{i\in I}$ ist aber nicht nur für dreielementige oder endliche Indexmengen $I$ sondern für beliebige Indexmengen $I$ definiert. Zum konkreten Hinschreiben beziehungsweise vorstellen tut man sich etwa bei

  $$\{0, 1\}^{\mathbb{R}} = \{(x_r)_{r\in\mathbb{R}} : x_r \in \{0, 1\}\}$$

  hart. Aber $(x_r)_{r\in\mathbb{R}} \in \{0, 1\}^{\mathbb{R}}$ besteht aus $x_r \in \{0, 1\}$, wobei die Indexmenge $\mathbb{R}$ ist, zB. ist $(x_r)_{r\in\mathbb{R}}$ mit $x_r = 1$ falls $r \in \mathbb{Q}$ und $x_r = 0$ falls $r \in \mathbb{R}\backslash\mathbb{Q}$ in $\{0, 1\}^{\mathbb{R}}$ enthalten.

**Lemma 1.1.10.** *Sind $I$ und $J$ Mengen und $X_i \subseteq X$ für $i \in I$ und $Y_j \subseteq Y$ für $j \in J$, dann gilt*

$$\left(\bigcup_{i\in I} X_i\right) \cap \left(\bigcup_{j\in J} Y_j\right) = \bigcup_{(i,j)\in I\times J} (X_i \cap Y_j) \tag{1.1.4}$$

$$\left(\bigcap_{i\in I} X_i\right) \cup \left(\bigcap_{j\in J} Y_j\right) = \bigcap_{(i,j)\in I\times J} (X_i \cup Y_j) \tag{1.1.5}$$

*Beweis.* Zum Beweis von (1.1.4):

„$\subseteq$“: Für $x \in \left(\bigcup_{i\in I} X_i\right) \cap \left(\bigcup_{j\in J} Y_j\right)$ gibt es $i_0 \in I$ und $j_0 \in J$ mit

$$x \in X_{i_0} \cap Y_{j_0} \subseteq \bigcup_{(i,j)\in I\times J} (X_i \cap Y_j).$$

„$\supseteq$“: Zu $x \in \bigcup_{(i,j)\in I\times J} (X_i \cap Y_j)$ gibt es $i_0 \in I$ und $j_0 \in J$ mit

$$x \in X_{i_0} \cap Y_{j_0} \subseteq \left(\bigcup_{i\in I} X_i\right) \cap \left(\bigcup_{j\in J} Y_j\right).$$

Der Beweis von (1.1.5) bleibt als Übungsaufgabe. □

**Beispiel$^{\#}$ 2.** Für endliche Indexmengen $I$ und $J$ werden (1.1.4) und (1.1.5) recht übersichtlich, zB ist $(X_1 \cup X_2) \cap (Y_1 \cup Y_2) = (X_1 \cap Y_1) \cup (X_1 \cap Y_2) \cup (X_2 \cap Y_1) \cup (X_2 \cap Y_2)$.

## 1.2 Elementare Aussagenlogik

Eine **Aussage** $A$ (zB. ein Satz oder eine Formel) hat immer genau einen der beiden Wahrheitsgehalte, nämlich entweder „WAHR“ oder „FALSCH“.

**Beispiel$^*$ 1.2.1.** [2]

- Ist $M$ eine Menge, $a$ ein Objekt, so ist „$a \in M$“ eine Aussage, denn entweder ist $a$ ein Element von $M$, dann ist „$a \in M$“ WAHR, oder $a$ ist kein Element von $M$, dann ist „$a \in M$“ FALSCH.
- „7 ist eine Primzahl“ ist eine Aussage (mit dem Wahrheitsgehalt WAHR)
- „$3 < 1$“ ist eine Aussage (mit dem Wahrheitsgehalt FALSCH)

[2] Bei einem Beispiel* kann ein “neues“ Konzept vorkommen, dem wir in der Vorlesung noch nicht begegnet sind – wie hier etwa “Primzahl”. Mit der Zeit sollten die Beispiel* verschwinden und alle in den Beispielen vorkommenden Begriffe auch in der Vorlesung definiert sein – aber zu Beginn der Vorlesung können wir nicht auf einen Schlag alles machen.

- „violett“ ist keine Aussage

Aus einer Aussage läßt sich eine neue Aussage basteln: Ist etwa $A$ eine Aussage, dann ist „A ist falsch“ wieder eine Aussage, die **Negation** von $A$, in Zeichen $\neg A$. In der **Wahrheitstafel** sieht man, wie WAHR und FALSCH von $A$ und $\neg A$ zusammenhängen:

| $A$ | W | F |
|---|---|---|
| $\neg A$ | F | W |

Hat man die Aussagen $A$ und $B$, so lassen sich beispielsweise folgende Verknüpfungen angeben:

| $A$ | $B$ | $A \wedge B$ | $A \vee B$ | $A \Rightarrow B$ |
|---|---|---|---|---|
| W | W | W | W | W |
| W | F | F | W | F |
| F | W | F | W | W |
| F | F | F | F | W |

Beachte: Die Verknüpfung $\wedge$ = „UND“ entspricht dem umgangssprachlichen Gebrauch. Das $\vee$ = „ODER“ ist ein „vel“ und kein „aut“, also auch wahr, wenn $A$ und $B$ wahr sind und nicht im Sinn von einem ausschließenden „entweder ... oder “ zu verstehen. Die Verknüpfung $A \Rightarrow B$ „WENN $A$ DANN $B$“ (oder auch „AUS $A$ FOLGT $B$“) ist auf den ersten Blick auch etwas gewöhnungsbedürftig, berücksichtigt aber, daß man aus einer falschen Aussage (zB. $1 = 2$) durchaus eine richtige Aussage folgern kann. (zB. nach Multiplikation mit 0 folgt: $0 = 0$)

**Lemma 1.2.2.** *Mit $\wedge$ und $\vee$ verknüpfte Aussagen $A$ und $B$ negiert man, indem man die Aussagen negiert und $\wedge$ mit $\vee$ vertauscht.*

*Beweis.* Wir schreiben uns die Wahrheitstafel für die möglichen Kombinationen auf

| $A$ | $B$ | $\neg(A \wedge B)$ | $(\neg A) \vee (\neg B)$ | $\neg(A \vee B)$ | $(\neg A) \wedge (\neg B)$ |
|---|---|---|---|---|---|
| W | W | F | F | F | F |
| W | F | W | W | F | F |
| F | W | W | W | F | F |
| F | F | W | W | W | W |

und sehen die Behauptung aus dem Vergleich der 3. und 4. beziehungsweise 5. und 6. Spalte der Wahrheitstafel. □

**Bemerkung 1.2.3.** Auch $\Rightarrow$ läßt sich durch $\neg$ und $\vee$ schreiben; die Wahrheitstafel

| $A$ | $B$ | $A \Rightarrow B$ | $(\neg A) \vee B$ |
|---|---|---|---|
| W | W | W | W |
| W | F | F | F |
| F | W | W | W |
| F | F | W | W |

zeigt das. Wir sehen daraus insbesondere, daß die Verneinung einer „WENN ... DANN“ Aussage keine „WENN ... DANN“ Aussage mehr ist, sondern $\neg(A \Rightarrow B)$ ist $A \wedge (\neg B)$. Die Kombination $A \Leftrightarrow B$, also „A GENAU DANN WENN B“ ist $(A \Rightarrow B) \wedge (B \Rightarrow A)$.

**Bemerkung 1.2.4.** Ist $I$ eine Menge und ist für jedes $i \in I$ eine Aussage $A(i)$ gegeben, so schreibt man

$\forall i \in I : A(i)$, falls für alle $i \in I$ die Aussage $A(i)$ richtig ist.

$\exists i \in I : A(i)$, falls es ein $i \in I$ gibt, so daß die Aussage $A(i)$ richtig ist.

und erhält so aus den Aussagen $(A(i))_{i\in I}$ diese beiden neuen Aussagen. Man bezeichnet $\forall$ und $\exists$ als **Quantoren**. Aus der Wahrheitstafel sieht man folgende Regeln:

$$\neg\Big(\forall i \in I : A(i)\Big) \text{ ist dieselbe Aussage wie } \exists i \in I : \Big(\neg A(i)\Big) \tag{1.2.1}$$

$$\neg\Big(\exists i \in I : A(i)\Big) \text{ ist dieselbe Aussage wie } \forall i \in I : \Big(\neg A(i)\Big) \tag{1.2.2}$$

## 1.3 Funktionen

**Definition 1.3.1.** *Es seien $X, Y$ Mengen; ein Tripel $f = (X, Y, \Gamma)$ heißt eine* **Funktion** *(oder* **Abbildung***) auf $X$ mit Werten in $Y$, wenn gilt:*

1. *$\Gamma \subseteq X \times Y$.*
2. *Zu jedem $x \in X$ gibt es genau ein $y \in Y$ mit $(x, y) \in \Gamma$.*

*In diesem Fall heißt $X$ der* **Definitionsbereich** *von $f$, $Y$ der* **Wertevorrat** *von $f$ und $\Gamma$ der* **Graph** *von $f$. Eine gebräuchliche Schreibweise für eine Funktion $f$ auf $X$ mit Werten in $Y$ ist $f : X \to Y$.*

**Beispiel* 1.3.2.**

- Durch $X := \{q \in \mathbb{Q} : -1 \leq q \leq 1\}$ und $\Gamma := \{(x, y) \in X \times X : x^2 + y^2 = 1\}$ wird keine Funktion $f = (X, X, \Gamma)$ definiert, da zB. $(0, 1), (0, -1) \in \Gamma$, also die Bedingung 2) an $\Gamma$ aus Definition 1.3.1 für $x = 0 \in X$ verletzt ist.

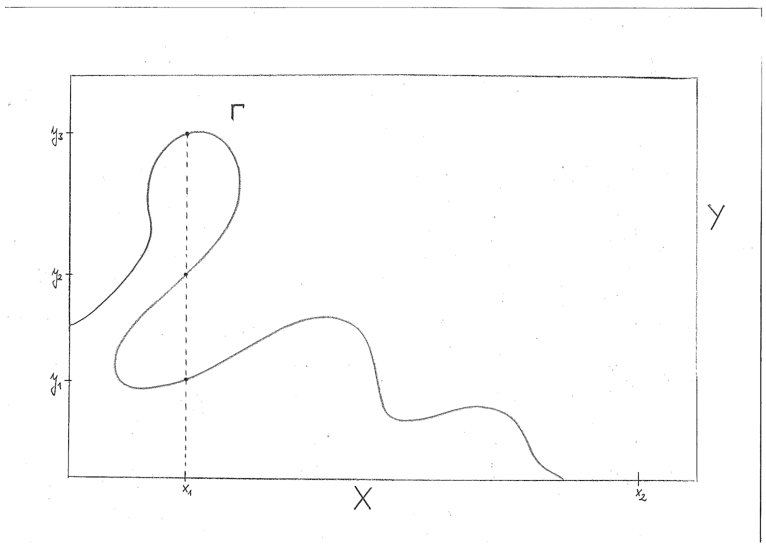
- Auch durch $X := \{q \in \mathbb{Q} : -1 \leq q \leq 1\}$ und $\Gamma := \{(x, 1) : x \geq 0\}$ wird keine Funktion $f = (X, X, \Gamma)$ definiert, da es zB. kein $y \in X$ mit $(-1, y) \in \Gamma$ gibt.

- Vor allem in der Physik wird eine Funktion oft durch $f(x) = ...$ definiert – das ist korrekt, wenn man noch den Definitionsbereich $X$ und den Wertevorrat $Y$ angibt und für jedes $x \in X$ durch $f(x)$ ein eindeutiges Element in $Y$ definiert wird. Eine Schreibweise, die diese Details berücksichtigt ist $\begin{array}{rccc} f: & X & \to & Y \\ & x & \mapsto & f(x) \end{array}$, zB. $\begin{array}{rccc} f: & \mathbb{Q} & \to & \mathbb{Q} \\ & x & \mapsto & x^2 \end{array}$.
  In der Schreibweise aus Definition 1.3.1 hätte man hier $f = (X, Y, \Gamma)$ mit $\Gamma = \{(x, f(x)) : x \in X\}$ bzw. im Beispiel $f = (\mathbb{Q}, \mathbb{Q}, \Gamma)$ mit $\Gamma = \{(x, x^2) : x \in \mathbb{Q}\}$.

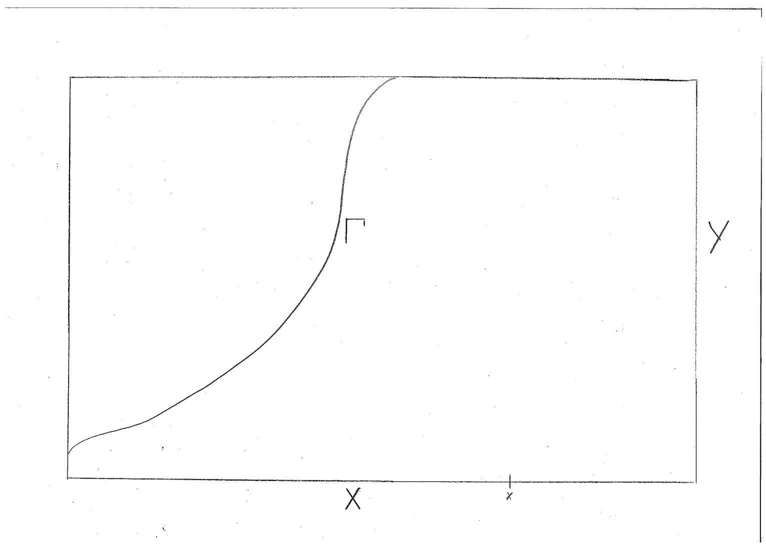
- Eine Aufgabe wie „bestimme den Definitionsbereich von $f(x) = \frac{1}{x-1}$ “ macht keinen Sinn!. Eine korrekte Aufgabe könnte so lauten: „Bestimme die größte Teilmenge $X \subseteq \mathbb{Q}$, so daß $\begin{array}{rccc} f: & X & \to & \mathbb{Q} \\ & x & \mapsto & \frac{1}{x-1} \end{array}$ eine Funktion definiert. Die Antwort auf diese Frage wäre dann $X = \mathbb{Q}\backslash\{1\}$, da man den Ausdruck $\frac{1}{x-1}$ für $x = 1$ nicht sinnvoll definieren

kann. Beachte aber, daß für alle $a \in \mathbb{Q}$ durch
$$\begin{array}{rcl} f_a : \mathbb{Q} & \to & \mathbb{Q} \\ x & \mapsto & \begin{cases} \frac{1}{x-1} & \text{für } x \neq 1 \\ a & \text{für } x = 1 \end{cases} \end{array}$$
eine Funktion definiert wird.

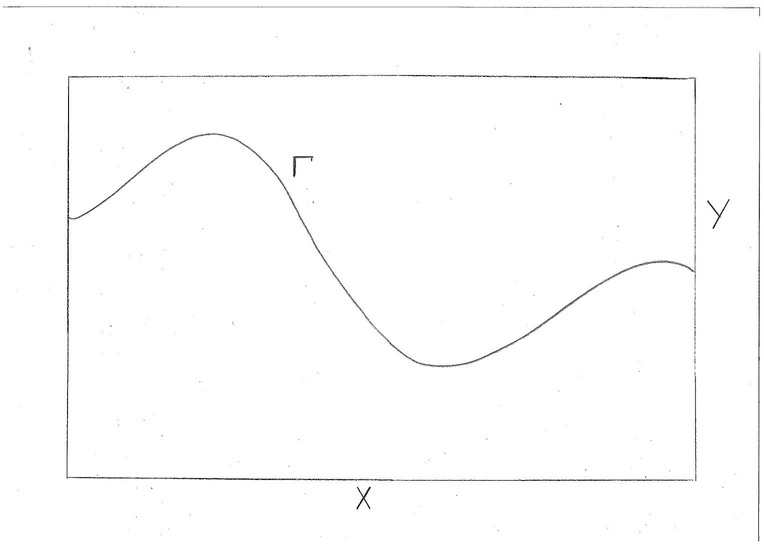
- Das im nächsten Bild skizzierte Tupel $(X, Y, \Gamma)$ ist keine Funktion mit Definitionsbereich $X$, denn zB. sind $(x_1, y_1), (x_1, y_2), (x_1, y_3) \in \Gamma$ oder es gibt kein $y \in Y$ mit $(x_2, y) \in \Gamma$.




- Das im nächsten Bild skizzierte Tupel $(X, Y, \Gamma)$ ist keine Funktion mit Definitionsbereich $X$, denn zB. gibt es kein $y \in Y$ mit $(x, y) \in \Gamma$.




- Das im nächsten Bild skizzierte Tupel $(X, Y, \Gamma)$ ist eine Funktion mit Definitionsbereich $X$




**Definition 1.3.3.** *Eine Funktion $f : X \to Y$ heißt genau dann*

- **injektiv**, *wenn für alle $w, x \in X$ gilt: Aus $f(w) = f(x)$ folgt $w = x$.*
- **surjektiv**, *wenn $f(X) := \{f(x) : x \in X\} = Y$ gilt.*
- **bijektiv**, *wenn $f$ injektiv und surjektiv ist.*

**Beispiel 1.3.4.**

a) Es sei $\emptyset \neq X$ eine Menge, dann ist $\begin{array}{rcl} \mathrm{id}_X : X & \to & X \\ x & \mapsto & x \end{array}$ eine bijektive Abbildung.

b) Es sei $\emptyset \neq X$ eine Menge, $Y \subseteq X$ mit $Y \neq X$, dann ist $\begin{array}{rcl} i : Y & \to & X \\ y & \mapsto & y \end{array}$ injektiv, aber nicht surjektiv.

c) Es sei $Y \subseteq X$, dann ist $\begin{array}{rcl} \mathbf{1}_Y : X & \mapsto & \{0, 1\} \\ x & \mapsto & \begin{cases} 1 & \text{für } x \in Y \\ 0 & \text{für } x \in X \setminus Y \end{cases} \end{array}$ nicht injektiv, sobald $Y$ oder $X \setminus Y$ mehr als ein Element besitzen und surjektiv, sobald $\emptyset \neq Y \neq X$ ist.

**Beispiel* 1.3.5.** Die verschiedensten Kombinationen der Eigenschaften nicht injektiv/ injektiv und nicht surjektiv / surjektiv sind möglich. Das geht sogar soweit, daß man für dieselbe Formel für den Funktionswert durch Wahl von Definitionsbereich und Wertevorrat zu den Beispielen kommt.

- $\begin{array}{rcl} f_1 : \mathbb{Z} & \to & \mathbb{Z} \\ x & \mapsto & x^2 \end{array}$ ist nicht injektiv da zB. $f(-1) = 1 = f(1)$ und nicht surjektiv, da $f_1(\mathbb{Z}) = \{x^2 : x \in \mathbb{Z}\} \subseteq \mathbb{N}_0 \neq \mathbb{Z}$.
- Für $X := \{n^2 : n \in \mathbb{Z}\}$ ist $\begin{array}{rcl} f_2 : \mathbb{Z} & \to & X \\ x & \mapsto & x^2 \end{array}$ nicht injektiv da zB. $f(-1) = 1 = f(1)$ und nach Definition von $X$ surjektiv.

- $\begin{array}{rcl} f_3 : \mathbb{N}_0 & \to & \mathbb{Z} \\ x & \mapsto & x^2 \end{array}$ ist injektiv – das folgt zB. da $(\mathbb{N}_0, \leq)$ total geordnet und $f_3$ streng monoton steigend ist – und nicht surjektiv, da zB. $-1 \notin X \subseteq \mathbb{Z}$ ist.

- $\begin{array}{rcl} f_4 : \mathbb{N}_0 & \to & X \\ x & \mapsto & x^2 \end{array}$ ist bijektiv – das folgt aus den Argumenten für $f_2$ und $f_3$.

**Lemma 1.3.6.** *Es seien $f : X \to Y$ und $g : Y \to Z$ Funktionen, dann definiert $g \circ f := (X, Z, \{(x, g(f(x))) : x \in X\})$, bzw. in der anderen Schreibweise $\begin{array}{rcl} g \circ f : X & \to & Z \\ x & \mapsto & g(f(x)) \end{array}$ eine Funktion auf $X$ mit Werten in $Z$. Sprechweise „$g$ nach $f$“ oder „$g$ komponiert mit $f$“ oder* **Komposition** *von $f$ mit $g$. Ist $h : Z \to W$ eine Funktion, so gilt:*

$$h \circ (g \circ f) = (h \circ g) \circ f. \tag{1.3.1}$$

*Beweis.* Offenbar ist $\Gamma := \{(x, g(f(x))) : x \in X\}$ der Graph einer Funktion auf $X$ mit Werten in $Z$, also die Komposition $g \circ f$ eine Funktion. Für $h : Z \to W$ gilt

$$h \circ (g \circ f) = (X, W, \{(x, h(g(f(x)))) : x \in X\}) = (h \circ g) \circ f.$$

□

**Beispiel* 1.3.7.** Vorsicht, die Reihenfolge läßt sich beim Komponieren von Funktionen nicht vertauschen: Sind $f : X \to Y$ und $g : Y \to Z$ Funktionen, so ist zwar $g \circ f : X \to Z$ definiert, aber $f \circ g$ existiert nur, wenn $g(Y) \subseteq X$ ist, und man $f(g(y))$ für alle $y \in Y$ bilden kann. Auch wenn $g \circ f$ und $f \circ g$ beide existieren, sind sie im Allgemeinen nicht gleich, zB. für $\begin{array}{rcl} f : \mathbb{Q} & \to & \mathbb{Q} \\ x & \mapsto & x+1 \end{array}$ und $\begin{array}{rcl} g : \mathbb{Q} & \to & \mathbb{Q} \\ x & \mapsto & x^2 \end{array}$ sind $\begin{array}{rcl} g \circ f : \mathbb{Q} & \to & \mathbb{Q} \\ x & \mapsto & (x+1)^2 \end{array}$ und $\begin{array}{rcl} f \circ g : \mathbb{Q} & \to & \mathbb{Q} \\ x & \mapsto & x^2+1 \end{array}$ verschieden.

**Definition 1.3.8.** *Es sei $f : X \to Y$ eine Funktion, dann sind auch*

$$\begin{array}{rcl} F : \mathcal{P}(X) & \to & \mathcal{P}(Y) \\ W & \mapsto & f(W) := \{f(w) : w \in W\} \end{array} \tag{1.3.2}$$

*und*

$$\begin{array}{rcl} f^{-1} : \mathcal{P}(Y) & \to & \mathcal{P}(X) \\ Z & \mapsto & f^{-1}(Z) := \{x \in X : f(x) \in Z\} \end{array} \tag{1.3.3}$$

*Funktionen. Zu $W \subseteq X$ heißt $f(W)$ das* **Bild** *von $W$ bei $f$ und zu $Z \subseteq Y$ heißt $f^{-1}(Z)$ das* **Urbild** *von $Z$ bei $f$.*

**Beispiel 1.3.9.** Ist $X = \{\square, \diamond, \heartsuit\}$ und $Y = \{\triangle, \bullet\}$ und $f : X \to Y$ definiert durch $f(\square) := \triangle$, $f(\diamond) := \bullet$ und $f(\heartsuit) := \triangle$. Dann ist $f^{-1}(\emptyset) = \emptyset$, $f^{-1}(\{\triangle\}) = \{\square, \heartsuit\}$ $f^{-1}(\{\bullet\}) = \{\diamond\}$ und $f^{-1}(Y) = X$. An diesem Beispiel sollte man sich auch noch einmal klar machen, daß das Urbild von $f : X \to Y$ eine Funktion $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ definiert. Das geht auch in diesem Beispiel, wo $f$ nicht injektiv ist – auch hier ist etwa $f^{-1}(\{\triangle\}) = \{\square, \heartsuit\} \in \mathcal{P}(X)$.

**Lemma 1.3.10.** *Es sei $f : X \to Y$ eine Funktion, $I$ eine Menge und $X_i \subseteq X$ und $Y_i \subseteq Y$ für alle $i \in I$. Dann gilt:*

*a)* $f^{-1}(\emptyset) = \emptyset$ *und* $f^{-1}(Y) = X$.

*b)* $f^{-1}\left(\bigcup_{i\in I} Y_i\right) = \bigcup_{i\in I} f^{-1}(Y_i)$

*c)* $f^{-1}\left(\bigcap_{i\in I} Y_i\right) = \bigcap_{i\in I} f^{-1}(Y_i)$

*d)* $f\left(\bigcup_{i\in I} X_i\right) = \bigcup_{i\in I} f(X_i)$

*e)* $f\left(\bigcap_{i\in I} X_i\right) \subseteq \bigcap_{i\in I} f(X_i)$

*Beweis.* von (a), (b) und (c):

a) Nach Definition der Funktion $f : X \to Y$ gibt es für jedes $x \in X$ einen eindeutigen Funktionswert $f(x) \in Y$, daher gilt nach Definition des Urbilds: $f^{-1}(Y) = \{x \in X : f(x) \in Y\} = X$ und $f^{-1}(\emptyset) = \{x \in X : f(x) \in \emptyset\} = \emptyset$.

b) „$\subseteq$“ Sei $x \in f^{-1}\left(\bigcup_{i\in I} Y_i\right)$, so gibt es ein $j \in I$ mit $f(x) \in Y_j$, dh. $x \in f^{-1}(Y_j) \subseteq \bigcup_{i\in I} f^{-1}(Y_i)$.

„$\supseteq$“ Sei $x \in \bigcup_{i\in I} f^{-1}(Y_i)$, so gibt es $j \in I$ mit $x \in f^{-1}(Y_j)$, dh. $f(x) \in Y_j \subseteq \bigcup_{i\in I} Y_i$, also $x \in f^{-1}\left(\bigcup_{i\in I} Y_i\right)$.

c) „$\subseteq$“ Für $x \in f^{-1}\left(\bigcap_{i\in I} Y_i\right)$ ist $f(x) \in \bigcap_{i\in I} Y_i$, dh. für alle $i \in I$ gilt: $f(x) \in Y_i$ oder $x \in f^{-1}(Y_i)$, also $x \in \bigcap_{i\in I} f^{-1}(Y_i)$.

„$\supseteq$“ Sei $x \in \bigcap_{i\in I} f^{-1}(Y_i)$, so ist $f(x) \in Y_i$ für alle $i \in I$, also $f(x) \in \bigcap_{i\in I} Y_i$ oder $x \in f^{-1}\left(\bigcap_{i\in I} Y_i\right)$.

$\square$

**Bemerkung# 1.** Eine Funktion $f : X \to Y$ ist genau dann injektiv, wenn $f^{-1}(f(V)) = V$ für alle $V \subseteq X$ gilt. Dies ist ferner äquivalent dazu daß in Lemma 1.3.10(e) Gleichheit gilt, also

$$f\left(\bigcap_{i\in I} X_i\right) = \bigcap_{i\in I} f(X_i)$$

für $X_i \subseteq X$ gilt. (Aufgabe!)

**Lemma 1.3.11.** *Es seien $f : X \to Y$ und $g : Y \to Z$ Funktionen.*

- *Sind $f$ und $g$ injektiv, dann ist auch $g \circ f$ injektiv.*
- *Sind $f$ und $g$ surjektiv, dann ist auch $g \circ f$ surjektiv.*
- *Sind $f$ und $g$ bijektiv, dann ist auch $g \circ f$ bijektiv.*

*Beweis.*

- Es seien $f$ und $g$ injektiv und $w, x \in X$ mit $(g \circ f)(w) = g(f(w)) = (g \circ f)(x) = g(f(x))$. Da nach Voraussetzung $g$ injektiv ist, folgt daraus $f(w) = f(x)$ und aus der Injektivität von $f$ folgt $w = x$, dh. $g \circ f$ ist injektiv.
- Es seien $f$ und $g$ surjektiv und $z \in Z$. Da $g$ surjektiv ist, gibt es ein $y \in Y$ mit $z = g(y)$ und weil $f$ surjektiv ist, gibt es ein $x \in X$ mit $y = f(x)$. Insgesamt ist also $(g \circ f)(x) = g(f(x)) = g(y) = z$ und damit $g \circ f$ surjektiv. $\square$

**Lemma 1.3.12.** *Es seien $X, Y \neq \emptyset$ und $f : X \to Y$ eine Funktion.*

*a) Es gibt genau dann eine Funktion $g : Y \to X$ mit $g \circ f = \mathrm{id}_X$, wenn $f$ injektiv ist.*

*b) Es gibt genau dann eine Funktion $g : Y \to X$ mit $f \circ g = \mathrm{id}_Y$, wenn $f$ surjektiv ist.*

*Beweis.*

a) „$\Rightarrow$“ Es sei $g : Y \to X$ eine Funktion mit $g \circ f = \mathrm{id}_X$ und $w, x \in X$ mit $f(x) = f(w)$. Dann ist $w = g(f(w)) = g(f(x)) = x$, also ist $f$ injektiv.

„$\Leftarrow$“ Es sei $f : X \to Y$ injektiv, dann ist für jedes $y \in Y$

$$f^{-1}(\{y\}) = \begin{cases} \emptyset & \text{falls } y \notin f(X) \\ \{x\} & \text{falls } y = f(x) \in f(X) \end{cases}$$

Wir wählen nun $x_0 \in X$ und definieren damit

$$\begin{aligned} g : Y &\to X \\ y &\mapsto \begin{cases} x & \text{falls } y = f(x) \in f(X) \\ x_0 & \text{falls } y \notin f(X) \end{cases} \end{aligned}$$

Damit ist $g$ auf ganz $Y$ definiert und für alle $x \in X$ gilt: $(g \circ f)(x) = g(f(x)) = x$.

b) „$\Rightarrow$“ Sei $g : Y \to X$ eine Funktion mit $f \circ g = \mathrm{id}_Y$, dann ist $Y = f(g(Y)) \subseteq f(X)$ und wegen $f(X) \subseteq Y$ folgt dann $Y = f(X)$, also ist $f$ surjektiv.

„$\Leftarrow$“ Es sei $f : X \to Y$ surjektiv, dann ist $f^{-1}(\{y\}) \neq \emptyset$ für alle $y \in Y$. Wir können daher für alle $y \in Y$ ein $g(y) \in f^{-1}(\{y\})$ auswählen und bekommen daher für die Funktion $\begin{aligned} g : Y &\to X \\ y &\mapsto g(y) \end{aligned}$ und für alle $y \in Y$ die Gleichheit $(f \circ g)(y) = f(g(y)) = y$. $\square$

**Satz 1.3.13.** *Es sei $\emptyset \neq X, Y$ und $f : X \to Y$ eine Funktion. Es gibt genau dann eine Funktion $g : Y \to X$ mit $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$, wenn $f$ bijektiv ist. In diesem Fall ist diese Funktion $g$ eindeutig und heißt die* **Umkehrfunktion** *von $f$; Schreibweise: $f^{-1} := g$. Für eine Funktion $h : Y \to X$ sind in diesem Fall äquivalent:*

*a) $h = f^{-1}$*

*b) $h \circ f = \mathrm{id}_X$*

*c) $f \circ h = \mathrm{id}_Y$*

*Beweis.* Nach Lemma 1.3.12 ist $f : X \to Y$ genau dann bijektiv, wenn es $g_1, g_2 : Y \to X$ mit $g_1 \circ f = \mathrm{id}_X$ und $f \circ g_2 = \mathrm{id}_Y$ gibt. Es sei nun $f$ bijektiv, dann tritt der Fall $y \notin f(X)$ aus dem Beweis von Lemma 1.3.12, a) „$\Leftarrow$“ nicht auf und $g_1 = g_2 =: g$ ist eindeutig. Beim Beweis der Äquivalenzen gilt offenbar a) $\Rightarrow$ b) und a) $\Rightarrow$ c), sowie b) und c) $\Rightarrow$ *a*).

b) $\Rightarrow$ c): Angenommen $h \circ f = \mathrm{id}_X$ und $f \circ h \neq \mathrm{id}_Y$ (was die Negation von b) $\Rightarrow$ c) ist!) , dann gilt $f = f \circ (h \circ f)$ und es gibt ein $y_0 \in Y$ mit $(f \circ h)(y_0) \neq y_0$. Da $f$ surjektiv ist, gibt es ein $x_0 \in X$ mit $y_0 = f(x_0)$, also ist $y_0 = f(x_0) = (f \circ (h \circ f))(x_0) = ((f \circ h) \circ f)(x_0) = (f \circ h)(y_0) \neq y_0$; Widerspruch.

c) $\Rightarrow$ b): Angenommen $f \circ h = \mathrm{id}_Y$ und $h \circ f \neq \mathrm{id}_X$, dann gilt $f = (f \circ h) \circ f$ und es gibt $x_0 \in X$ mit $x_1 := (h \circ f)(x_0) \neq x_0$. Daraus folgt $f(x_0) = ((f \circ h) \circ f)(x_0) = f((h \circ f)(x_0)) = f(x_1)$, was aber nicht sein kann, da $f$ injektiv ist. $\square$

**Bemerkung 1.3.14.** Zwischen der Umkehrfunktion $f^{-1}$ und dem Urbild $f^{-1}$ einer Funktion $f : X \to Y$ sollte es zu keinen Verwechslungen kommen, wenn man einige Dinge beachtet:

- $\mathcal{P}(Y)$ ist der Definitionsbereich für das Urbild $f^{-1}$, also bedeutet $f^{-1}(Z)$ für $Z \subseteq Y$ immer das Urbild von $Z$ unter $f$ – egal ob es eine Umkehrfunktion gibt oder nicht.
- Nur wenn $f$ bijektiv ist, gibt es überhaupt eine Umkehrfunktion $f^{-1} : Y \to X$, daher bezeichnet also in diesem Fall $f^{-1}(y)$ für $y \in Y$ immer den Funktionswert der Umkehrfunktion an der Stelle $y$.
- Für bijektives $f : X \to Y$ und einelementige Teilmengen $\{y\} \subseteq Y$ erhält man noch
$$f^{-1}(\{y\}) = \{f^{-1}(y)\}. \tag{1.3.4}$$
Bei bijektivem $f : X \to Y$ gibt es für $Z \subseteq Y$ auch bei $f^{-1}(Z)$ keine Verwechslungsgefahr; $f^{-1}(Z)$ ist sowohl das Bild von $Z$ bei der Umkehrabbildung $f^{-1} : Y \to X$ als auch das Urbild von $Z$ bei $f$.

**Beispiel* 1.3.15.** Die Bedingung, daß $f : X \to Y$ bijektiv ist, ist nach dem obigen Satz äquivalent zur Existenz einer Umkehrfunktion $f^{-1} : Y \to X$. Wie paßt das mit den inversen trigonometrischen Funktionen zusammen? Etwa $\begin{array}{rcl}\sin : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \sin(x)\end{array}$ ist weder injektiv (zB. $\sin(0) = \sin(\pi) = 0$) noch surjektiv (zB. $\sin(\mathbb{R}) = [-1,1]$). Aber $\begin{array}{rcl}f : [-\frac{\pi}{2}, \frac{\pi}{2}] & \to & [-1,1] \\ x & \mapsto & \sin(x)\end{array}$ ist bijektiv und $\begin{array}{rcl}\arcsin : [-1,1] & \to & [-\frac{\pi}{2}, \frac{\pi}{2}] \\ y & \mapsto & \arcsin(y)\end{array}$ ist nicht die Umkehrfunktion von sin (denn die gibt es nicht!), aber arcsin ist die Umkehrfunktion von $f$.

**Bemerkung# 2.**
Wenn man in der Definition 1.1.8 des cartesischen Produkts $I-$Tupel vermeiden möchte, dann kann man zuerst das cartesische Produkt von zwei Mengen und damit Funktionen definieren und dann wird das cartesische Produkt $\prod\limits_{i \in I} X_i$ als

$$X := \{f : I \to \bigcup_{i \in I} X_i : f \text{ ist Funktion und für jedes } i \in I \text{ gilt: } f(i) \in X_i\}$$

definiert, genauer gesagt ist die Abbildung

$$\begin{array}{rcl}\Phi : X & \to & \prod\limits_{i \in I} X_i \\ f & \mapsto & (f(i))_{i \in I}\end{array}$$

bijektiv (Aufgabe!) und mittels $\Phi$ läßt sich $X$ und $\prod\limits_{i \in I} X_i$ miteinander identifizieren.

## 1.4 Äquivalenz- und Ordnungsrelationen

**Definition 1.4.1.** *Es sei $X$ eine Menge und $R \subseteq X \times X$ und wir schreiben $x \sim y$, wenn $(x, y) \in R$. Dann heißt $\sim$ eine* **Äquivalenzrelation** *auf $X$, wenn gilt:*

- **Reflexivität**: *Für alle $x \in X$ gilt $x \sim x$.*
- **Symmetrie**: *Für alle $x, y \in X$ gilt: Aus $x \sim y$ folgt $y \sim x$.*
- **Transitivität**: *Für alle $x, y, z \in X$ gilt: Aus $x \sim y$ und $y \sim z$ folgt $x \sim z$.*

*Ist $\sim$ eine Äquivalenzrelation auf $X$ und $x \in X$, so heißt*

$$[x]_\sim := \{y \in X : y \sim x\}$$

*die* **Äquivalenzklasse** *von $x$ bezüglich $\sim$ und*

$$X/_\sim := \{[x]_\sim : x \in X\}$$

*die Menge der Äquivalenzklassen bezüglich $\sim$.*

**Beispiel 1.4.2.**

- Ist $\emptyset \neq X$ eine Menge, so wird durch die Gleichheit $=$ eine Äquivalenzrelation auf $X$ gegeben; es gilt $[x]_= = \{x\}$ für $x \in X$.
- Betrachte nocheinmal das Beispiel 1.1.1: Dann kann man auf $M = \{1, 2, 3, 4, 5, 6\}^2$ durch

  $$(x_1, y_1) \sim (x_2, y_2) \quad :\Leftrightarrow \quad (x_1 = x_2 \text{ und } y_1 = y_2) \text{ oder } (x_1 = y_2 \text{ und } x_2 = y_1)$$

  eine Äquivalenzrelation $\sim$ auf $M$ definieren und mit der Notation aus Beispiel 1.1.1 gilt $[(x, y)]_\sim = \langle x; y\rangle$ und $M/_\sim = N$.

**Definition 1.4.3.** *Es seien $X$ und $I$ nichtleere Mengen und $X_i \subseteq X$ für alle $i \in I$. Dann bilden $X_i, i \in I$ eine* **Zerlegung** *von $X$, wenn gilt:*

*a)* $X_i \neq \emptyset$ *für alle $i \in I$.*

*b)* $\bigcup_{i \in I} X_i = X$

*c) Die Mengen $X_i, i \in I$ sind* **paarweise disjunkt**, *dh. für alle $i, j \in I$ mit $i \neq j$ gilt $X_i \cap X_j = \emptyset$.*

**Satz 1.4.4.** *Es sei $\sim$ eine Äquivalenzrelation auf $X$, dann bilden die Äquivalenzklassen $[x]_\sim, [x]_\sim \in X/_\sim$ eine Zerlegung von $X$. Umgekehrt definiert jede Zerlegung $X_i, i \in I$ von $X$ durch*

$$x \simeq y \quad :\Leftrightarrow \quad \text{Es gibt } i \in I \text{ mit } x, y \in X_i \tag{1.4.1}$$

*eine Äquivalenzrelation $\simeq$ auf $X$.*

*Beweis.* Es sei $\sim$ eine Äquivalenzrelation auf $X$, dann gilt wegen $x \sim x$, daß $[x]_\sim \neq \emptyset$ und $\bigcup_{[x]_\sim \in X/\sim} [x]_\sim = X$. Sind nun $x, y \in X$ mit $[x]_\sim \cap [y]_\sim \neq \emptyset$, so gibt es $z \in X$ mit $x \sim z$ und $z \sim y$. Da $\sim$ transitiv und symmetrisch ist, gilt auch $x \sim y$ und $w \sim y$ für alle $w \in [x]_\sim$ bzw. $v \sim x$ für alle $v \in [y]_\sim$, also gilt $[x]_\sim = [y]_\sim$.

Ist $X_i, i \in I$ eine Zerlegung von $X$ und $\simeq$ wie in (1.4.1) definiert, dann gilt:

- Für alle $x \in X$ gibt es (genau ein) $i \in I$ mit $x \in X_i$ und daher ist $x \simeq x$.
- Ist $x \simeq y$, so gibt es ein $i \in I$ mit $x, y \in X_i$ und daher ist auch $y \simeq x$.
- Sind $x \simeq y$ und $y \simeq z$, so gibt es $j \in I$ mit $x, y \in X_j$. Da die Mengen $X_i, i \in I$ paarweise disjunkt sind folgt dann aus $y \simeq z$ auch $y, z \in X_j$, also gilt auch $x \simeq z$.

□

**Beispiel* 1.4.5.** Auf $\mathbb{N}_0$ führen wir Division durch 3 mit Rest aus und bekommen eine Zerlegung von $\mathbb{N}_0$ in die paarweise disjunkten Mengen $X_0 := \{n = 3 \cdot k : k \in \mathbb{N}_0\}$, $X_1 := \{n = 3 \cdot k + 1 : k \in \mathbb{N}_0\}$ und $X_2 := \{n = 3 \cdot k + 2 : k \in \mathbb{N}_0\}$. Genausogut könnten wir nach Satz 1.4.4 auf $\mathbb{N}_0$ durch $m \sim n$ genau dann wenn $m$ und $n$ beim Teilen durch 3 den gleichen Rest lassen, eine Äquivalenzrelation definieren. Dann ist etwa $X_1$ die Äquivalenzklasse von 1, 4 etc.

Dieses Beispiel läßt sich weiter verallgemeinern: Statt die Division durch 3 mit Rest zu betrachten, können wir ebenso die Division durch eine natürliche Zahl $n$ mit Rest betrachten und erhalten dann eben $n$ Äquivalenzklassen.

**Definition 1.4.6.** *Es sei $X$ eine nichtleere Menge, $R \subseteq X \times X$; schreibe $x \leq y$ falls $(x, y) \in R$. Dann heißt $\leq$ eine* **Ordnungsrelation** *auf $X$ und $(X, \leq)$ eine* **geordnete Menge***, wenn gilt:*

- **Reflexivität:** *$x \leq x$ für alle $x \in X$.*
- **Antisymmetrie:** *Aus $x \leq y$ und $y \leq x$ folgt $x = y$.*
- **Transitivität:** *Aus $x \leq y$ und $y \leq z$ folgt $x \leq z$.*

*Für $x_1, x_2 \in X$ schreibt man $x_1 < x_2$ als Abkürzung von $x_1 \leq x_2$ und $x_1 \neq x_2$. $(X, \leq)$ heißt* **total geordnet***, wenn für alle $x, y \in X$ gilt: $x \leq y$ oder $y \leq x$.*

**Definition 1.4.7.** *Es sei $(X, \leq)$ eine geordnete Menge und $Y \subseteq X$, dann heißt*

- *$x \in X$ eine* **obere Schranke** *von $Y$ in $(X, \leq)$, wenn $y \leq x$ für alle $y \in Y$ erfüllt ist.*
- *$x \in X$ eine* **untere Schranke** *von $Y$ in $(X, \leq)$, wenn $x \leq y$ für alle $y \in Y$ erfüllt ist.*
- *$x \in X$ heißt* **Maximum** *von $Y$ in $(X, \leq)$, wenn $x$ obere Schranke von $Y$ in $(X, \leq)$ ist und $x \in Y$ gilt.*

- *$x \in X$ heißt* **Minimum** *von $Y$ in $(X, \leq)$, wenn $x$ untere Schranke von $Y$ in $(X, \leq)$ ist und $x \in Y$ gilt.*

- *$x \in X$ heißt* **Supremum** *von $Y$ in $(X, \leq)$, wenn $x$ obere Schranke von $Y$ ist und für jede weitere obere Schranke $z$ von $Y$ in $(X, \leq)$ gilt: $x \leq z$*

- *$x \in X$ heißt* **Infimum** *von $Y$ in $(X, \leq)$, wenn $x$ untere Schranke von $Y$ ist und für jede weitere untere Schranke $z$ von $Y$ in $(X, \leq)$ gilt: $z \leq x$*

- *$(X, \leq)$ heißt* **induktiv geordnet***, wenn jede total geordnete Teilmenge $Y \subseteq X$ eine obere Schranke in $X$ besitzt.*

- *$x \in X$ heißt ein* **maximales Element** *von $X$, wenn es kein $y \in X$ mit $x \leq y$ und $y \neq x$ gibt.*

**Beispiel 1.4.8.**

- Ist $\emptyset \neq Z$, so wird $X := \mathcal{P}(Z)$ durch $\subseteq$ zu einer geordneten Menge $(\mathcal{P}(Z), \subseteq)$

- Das letzte Beispiel zeigt, daß nicht jede Ordnung total sein muß. Zum Beispiel für $Z = \{1, 2\}$ lassen sich $Z_1 = \{1\} \in \mathcal{P}(Z)$ und $Z_2 = \{2\} \in \mathcal{P}(Z)$ nicht miteinander vergleichen.

- Nichtsdestotrotz ist für jede Menge $Z$ die Potenzmenge $(\mathcal{P}(Z), \subseteq)$ induktiv geordnet, denn $Z$ ist eine obere Schranke für jede (total geordnete) Teilmenge $Y \subseteq \mathcal{P}(Z)$.

- Ein maximales Element braucht nicht eindeutig zu sein, wie $Z = \{1, 2\}$ und $X = \mathcal{P}(Z)\setminus\{Z\} = \{\emptyset, \{1\}, \{2\}\}$ zeigt, wo $\{1\}$ und $\{2\}$ maximale Elemente (bzgl. $\subseteq$) sind. Ist $(X, \leq)$ totalgeordnet und $x \in X$ ein maximales Element, so ist $x$ eindeutig bestimmt – Aufgabe!

**Lemma 1.4.9.** *Es sei $(X, \leq)$ eine geordnete Menge und $Y \subseteq X$.*

- *Existiert ein Supremum $x$ von $Y$ in $(X, \leq)$, so ist es eindeutig bestimmt; Schreibweise:* $\sup(Y)$.

- *Existiert ein Infimum $x$ von $Y$ in $(X, \leq)$, so ist es eindeutig bestimmt; Schreibweise:* $\inf(Y)$.

- *Existiert ein Maximum $x$ von $Y$ in $(X, \leq)$, so ist es eindeutig bestimmt; Schreibweise:* $\max(Y)$. *In diesem Fall existiert auch $sup(Y)$ und es gilt* $\max(Y) = \sup(Y)$.

- *Existiert ein Minimum $x$ von $Y$ in $(X, \leq)$, so ist es eindeutig bestimmt; Schreibweise:* $\min(Y)$. *In diesem Fall existiert auch $inf(Y)$ und es gilt* $\min(Y) = \inf(Y)$.

*Beweis.*

- Sind $x, x'$ ein Supremum von $Y$, so sind $x$ und $x'$ obere Schranken von $Y$. Dann folgt $x \leq x'$, da $x$ Supremum und $x'$ obere Schranke ist und es folgt ebenso $x' \leq x$, da auch $x'$ Supremum und $x$ obere Schranke ist.

- Sind $x, x'$ ein Maximum von $Y$ so sind $x, x' \in Y$ obere Schranken von $Y$. Dann folgt $x \leq x'$, da $x'$ obere Schranke und $x \in Y$ ist; ebenso $x' \leq x$.

  Existiert also $\max(Y)$, dann gilt $\max(Y) \leq z$ für jede obere Schranke $z$ von $Y$ – nach Definition der oberen Schranke, da $\max(Y) \in Y$. Da $\max(Y)$ eine obere Schranke von $Y$ ist, folgt damit $\sup(Y) = \max(Y)$ aus der Definition von $\sup(Y)$. □

**Beispiel* 1.4.10.** Die $\leq$ Relation auf $\mathbb{R}$ ist ein Beispiel für eine totale Ordnung, aber $(\mathbb{R}, \leq)$ ist nicht induktiv geordnet, zB. hat $[0, \infty[$ keine obere Schranke – also auch kein Maximum und kein Supremum. Es ist $\min([0, \infty[) = \inf([0, \infty[) = 0$. Für $Y := \{x \in \mathbb{R} : -1 < x < 1\}$ ist $-1 = \inf(Y)$, $1 = \sup(Y)$ und $\min(Y)$ und $\max(Y)$ existieren nicht. Es gibt aber auch noch Beispiele für Mengen, die zwar obere und untere Schranken, aber kein Infimum oder Supremum besitzen, etwa $\{x \in \mathbb{Q} : x^2 < 2\}$ in $(\mathbb{Q}, \leq)$.

**Definition 1.4.11.** *Es seien $(X, \leq)$ und $(Y, \leq)$ geordnete Mengen und $f : X \to Y$ eine Funktion. Dann heißt*

- $f$ **monoton steigend**, *wenn für alle $x_1, x_2 \in X$ aus $x_1 \leq x_2$ folgt: $f(x_1) \leq f(x_2)$.*
- $f$ **streng monoton steigend**, *wenn für alle $x_1, x_2 \in X$ aus $x_1 < x_2$ folgt: $f(x_1) < f(x_2)$.*
- $f$ **monoton fallend**, *wenn für alle $x_1, x_2 \in X$ aus $x_1 \leq x_2$ folgt: $f(x_2) \leq f(x_1)$.*
- $f$ **streng monoton fallend**, *wenn für alle $x_1, x_2 \in X$ aus $x_1 < x_2$ folgt: $f(x_2) < f(x_1)$.*

**Beispiel 1.4.12.** Ist $X$ eine Menge, dann ist $\begin{array}{ccc} \mathcal{P}(X) & \to & \mathcal{P}(X) \\ Y & \mapsto & X \backslash Y \end{array}$ monoton fallend – sofern man $\mathcal{P}(X)$ durch $\subseteq$ ordnet. Für jede Abbildung $f : X \to Y$ ist $\begin{array}{rccc} F : & \mathcal{P}(X) & \to & \mathcal{P}(Y) \\ & W & \mapsto & f(W) \end{array}$ und $\begin{array}{rccc} f^{-1} : & \mathcal{P}(Y) & \to & \mathcal{P}(X) \\ & Z & \mapsto & f^{-1}(Z) \end{array}$ monoton steigend.

**Satz 1.4.13 (Zornsches Lemma).** *Jede induktiv geordnete Menge $(X, \leq)$ besitzt (mindestens) ein maximales Element.*

# Kapitel 2

# Die natürlichen Zahlen $\mathbb{N}$, vollständige Induktion und Mächtigkeit von Mengen

## 2.1 Die natürlichen Zahlen $\mathbb{N}$ und vollständige Induktion

**Definition 2.1.1.** *Die Menge $\mathbb{N}$ der natürlichen Zahlen erfüllt die* **Peano Axiome**

**P1:** *Es gibt ein ausgezeichnetes Element $1 \in \mathbb{N}$.*

**P2:** *Es gibt eine injektive Abbildung $N : \mathbb{N} \to \mathbb{N}\backslash\{1\}$.*

**P3:** *Ist $X \subseteq \mathbb{N}$ mit $1 \in X$ und $N(X) \subseteq X$, so gilt $X = \mathbb{N}$.*

**Bemerkung 2.1.2.** In der obigen Definition ist $N(n) =: n + 1$ der Nachfolger von $n$. Die Forderung, daß $N$ injektiv ist, stellt sicher, daß man beim Zählen 1,2,3,4,... nicht mehrmals auf die gleiche Zahl kommen kann. Das Peano Axiom (P3) hat die folgende – zunächst scheinbar völlig abstrakte –Konsequenz.

**Satz 2.1.3 (Vollständige Induktion).** *Für alle $n \in \mathbb{N}$ sei eine Aussage $A(n)$ gegeben und es gelte ferner:*

*a) $A(1)$ ist wahr.*

*b) Für jedes $n \in \mathbb{N}$ gilt: $A(n)$ wahr $\Rightarrow A(n + 1)$ wahr.*

*Dann ist $A(n)$ wahr für alle $n \in \mathbb{N}$.*

*Beweis.* Nach Voraussetzung a) ist $1 \in \{n \in \mathbb{N} : A(n) \text{ ist wahr}\} =: X$ und nach Voraussetzung b) folgt aus $n \in X$ auch $N(n) = n + 1 \in X$, dh. $N(X) \subseteq X$. Damit folgt $X = \mathbb{N}$ nach P3. □

Dieser Satz liefert jedoch den Bauplan für eine neue Beweismethode:

**Beispiel* 2.1.4.** Es sei $x \in \mathbb{R}, x \neq 1$, dann gilt für alle $n \in \mathbb{N}$:

$$1 + x + x^2 + ... + x^n = \frac{1 - x^{n+1}}{1 - x}. \tag{2.1.1}$$

Zum Beweis dieser Behauptung betrachten wir als $A(n)$ die Aussage

$$1 + x + x^2 + ... + x^n = \frac{1 - x^{n+1}}{1 - x}$$

und zeigen:

- **Induktionsanfang:** $1 + x = \dfrac{(1 + x)(1 - x)}{1 - x} = \dfrac{1 - x^2}{1 - x}$, dh. $A(1)$ ist wahr.

- **Induktionsschritt:** Es sei $A(n)$ wahr, dh. $1 + x + x^2 + ... + x^n = \dfrac{1 - x^{n+1}}{1 - x}$, dann gilt:

  $$1+x+x^2+...+x^n+x^{n+1} = \frac{1 - x^{n+1}}{1 - x}+x^{n+1} = \frac{1 - x^{n+1} + x^{n+1}(1 - x)}{1 - x} = \frac{1 - x^{n+2}}{1 - x},$$

  dh. dann ist auch $A(n + 1)$ wahr.

Somit sind die Voraussetzungen von Satz 2.1.3 erfüllt und daher gilt (2.1.1) für alle $n \in \mathbb{N}$.

**Beispiel 2.1.5.** Für alle $n \in \mathbb{N}$ ist $1 + 2 + ... + n = \dfrac{n(n + 1)}{2}$.
Zum Beweis nehmen wir $1 + 2 + ... + n = \frac{n(n+1)}{2}$ als Aussage $A(n)$ und zeigen:

- **Induktionsanfang:** $1 = \frac{1 \cdot 2}{2}$, also ist $A(1)$ wahr.

- **Induktionsschritt:** Es sei $A(n)$ wahr, also $1 + 2 + ... + n = \frac{n(n+1)}{2}$, dann ist

  $$1+2+...+n+(n+1) = \frac{n(n + 1)}{2} + (n+1) = (n+1)(\frac{n}{2}+1) = \frac{(n + 1)(n + 2)}{2},$$

  dh. $A(n + 1)$ ist wahr.

Somit sind die Voraussetzungen von Satz 2.1.3 erfüllt und daher gilt $1+2+...+n = \frac{n(n+1)}{2}$ für alle $n \in \mathbb{N}$.

**Definition 2.1.6.** *Es sei $Y$ eine Menge, dann heißt eine Funktion* $x : \begin{array}{ccc} \mathbb{N} & \to & Y \\ n & \mapsto & x_n \end{array}$ *eine* **Folge** *in $Y$. Schreibweise: $(x_n)_{n\in\mathbb{N}}$.*

**Warnung 2.1.7.** $(x_n)_{n\in\mathbb{N}}$ steht für eine Funktion mit Definitionsbereich $\mathbb{N}$. Also nicht die Folge $(x_n)_{n\in\mathbb{N}}$ mit der Menge der Folgenglieder $\{x_n : n \in \mathbb{N}\}$ – also in der obigen Schreibweise dem Bild $x(\mathbb{N})$ der Funktion $x = (x_n)_{n\in\mathbb{N}}$ verwechseln. Zum Beispiel ist

$$((-1)^n)_{n\in\mathbb{N}} = \begin{pmatrix} \mathbb{N} & \to & \{-1, 1\} \\ n & \mapsto & (-1)^n \end{pmatrix}$$

aber $\{(-1)^n : n \in \mathbb{N}\} = \{-1, 1\}$

**Satz 2.1.8** (Rekursionssatz). [1] *Es sei $Y$ eine Menge, $y \in Y$ und $f : Y \to Y$ eine Funktion. Dann gibt es genau eine Funktion $\varphi : \mathbb{N} \to Y$ mit $\varphi(1) = y$ und $\varphi \circ N = f \circ \varphi$. In der Schreibweise mit Folgen bedeutet das: Es gibt genau eine Folge $(\varphi_n)_{n\in\mathbb{N}}$ in $Y$ mit $\varphi_1 = y$ und $\varphi_{n+1} = f(\varphi_n)$ für alle $n \in \mathbb{N}$.*

*Beweis.* Wir zeigen die **Eindeutigkeit** von $\varphi$: Es seien $\varphi_1 : \mathbb{N} \to Y$ und $\varphi_2 : \mathbb{N} \to Y$ Funktionen mit $\varphi_1(1) = \varphi_2(1) = y$, $\varphi_1 \circ N = f \circ \varphi_1$ und $\varphi_2 \circ N = f \circ \varphi_2$. Dann gilt $\varphi_1(n) = \varphi_2(n)$ für alle $n \in \mathbb{N}$, wie wir durch vollständige Induktion zeigen (und damit $\varphi_1 = \varphi_2$):

- Induktionsanfang $n = 1$: Es gilt nach Voraussetzung $\varphi_1(1) = y = \varphi_2(1)$.
- Induktionssschluß $n \to n+1$: Nach Induktionsvoraussetzung ist $\varphi_1(n) = \varphi_2(n)$, also

$$\begin{aligned} \varphi_1(n+1) &= \varphi_1(N(n)) = (\varphi_1 \circ N)(n) = f(\varphi_1(n)) = f(\varphi_2(n)) \\ &= (\varphi_2 \circ N)(n) = \varphi_2(n+1). \end{aligned}$$

Damit gilt $\varphi_1 = \varphi_2$ und die gesuchte Funktion $\varphi$ ist eindeutig – sofern sie existiert.

Zur **Existenz** von $\varphi$ betrachte

$$\mathcal{H} := \{H \subseteq \mathbb{N} \times Y : (1, y) \in H \text{ und } (n, x) \in H \Rightarrow (N(n), f(x)) \in H\},$$

dann ist $\mathbb{N} \times Y \in \mathcal{H}$ und für alle $H_i, i \in I$ mit $H_i \in \mathcal{H}$ gilt $\bigcap\limits_{i\in I} H_i \in \mathcal{H}$, insbesondere gilt $D := \bigcap\limits_{H\in\mathcal{H}} H \in \mathcal{H}$. Wir zeigen nun – wieder durch Induktion, daß $D$ der Graph einer Funktion $\varphi : \mathbb{N} \to Y$ ist:

- Induktionsanfang n=1: $(1, y) \in H$ für jedes $H \in \mathcal{H}$, daher gilt $(1, y) \in D = \bigcap\limits_{H\in\mathcal{H}} H$.
  Angenommen es gibt $z \in Y$ mit $y \neq z$ und $(1, z) \in D$, so ist auch $D\backslash\{(1, z)\} \in \mathcal{H}$, woraus der Widerspruch $D\backslash\{(1, z)\} \subsetneq D = \bigcap\limits_{H\in\mathcal{H}} H \subseteq D\backslash\{(1, z)\}$ folgt.
- Induktionsschluß $n \to n + 1$: Nach Induktionsvoraussetzung gibt es genau ein $\varphi(n)$ mit $(n, \varphi(n)) \in D$. Nach Definition von $\mathcal{H}$ und $D$ folgt dann $(n + 1, f(\varphi(n))) = (N(n), f(\varphi(n))) \in D$. Angenommen es gibt $z \in Y$ mit $z \neq f(\varphi(n))$ und $(n + 1, z) \in D$, dann ist wieder $D\backslash\{(n + 1, z)\} \in \mathcal{H}$, also folgt wie eben der Widerspruch $D\backslash\{(n + 1, z)\} \subsetneq D \subseteq D\backslash\{(n + 1, z)\}$.

---

[1] Man findet manchmal andere Versionen des Rekursionssatzes, wie etwa

**Beispiel**$^{\#}$ **3** (Rekursionssatz 2. Variante). Es sei $Y$ eine Menge, $y \in Y$ und $f_n : Y^n \to Y$ eine Folge von Funktionen. Dann gibt es genau eine Folge $(g_n)_{n\in\mathbb{N}}$ in $Y$ mit $g_1 = y$ und $g_{n+1} = f_n(g_1, ..., g_n)$ für alle $n \in \mathbb{N}$.

mit denen man einige Beispiele leichter hinschreiben kann: Die Folge der **Fibonacci-Zahlen** bekommt man aus der 2. Variante des Rekursionssatzes durch die Wahl von $y = 1$ und $f_1 = \mathrm{id}_{\mathbb{N}}$ und $\begin{array}{rcl} f_n : \mathbb{N}^n & \to & \mathbb{N} \\ (y_1, ..., y_n) & \mapsto & y_{n-1} + y_n \end{array}$ für $n \geq 2$ und das ergibt die Folge $1, 1, 2, 3, 5, 8, 13, 21, ...$

Damit ist $D = \{(n, \varphi(n)) : n \in \mathbb{N}\}$ der Graph einer Funktion $\varphi : \mathbb{N} \to Y$ und es gilt $\varphi(1) = y$ (siehe Induktionsanfang) und die Eigenschaft $(N(n), f(\varphi(n))) \in D$ für $n \in \mathbb{N}$ bedeutet – da nun $D$ der Graph von $\varphi : \mathbb{N} \to Y$ ist – $(n+1, f(\varphi(n))) = (n+1, \varphi(n+1)) = (n+1, (\varphi \circ N)(n))$ für alle $n \in \mathbb{N}$ oder $f \circ \varphi = \varphi \circ N$. $\square$

Die wichtige Konsequenz, daß die Peano Axiome die natürlichen Zahlen eindeutig[2] festlegen, folgt aus dem Rekursionssatz:

**Lemma 2.1.9** (Eindeutigkeitssatz)**.** *Es sei $\tilde{\mathbb{N}}$ eine Menge mit $\tilde{1} \in \tilde{\mathbb{N}}$ und $\tilde{N} : \tilde{\mathbb{N}} \to \tilde{\mathbb{N}} \backslash \{\tilde{1}\}$ sei eine Funktion, die die Peano Axiome (P1),(P2) und (P3) erfüllt. Dann gibt es eine bijektive Abbildung $\phi : \mathbb{N} \to \tilde{\mathbb{N}}$ mit $\phi(1) = \tilde{1}$ und $\tilde{N} \circ \phi = \phi \circ N$.*

*Beweis.* Nach Satz 2.1.8 angewandt auf $Y = \tilde{\mathbb{N}}$, $y = \tilde{1}$ und $f = \tilde{N}$ gibt es genau eine Abbildung $\phi : \mathbb{N} \to \tilde{\mathbb{N}}$ mit $\phi(1) = \tilde{1}$ und $\phi \circ N = \tilde{N} \circ \phi$. Durch Vertauschen der Rollen von $\mathbb{N}$ und $\tilde{\mathbb{N}}$ findet man ein $\psi : \tilde{\mathbb{N}} \to \mathbb{N}$ mit $\psi(\tilde{1}) = 1$ und $\psi \circ \tilde{N} = N \circ \psi$. Sowohl für $\varphi = \mathrm{id}_{\mathbb{N}}$ als auch für $\varphi = \psi \circ \phi$ gilt $\varphi(1) = 1$ und $\varphi \circ N = N \circ \varphi$, also muß nach der Eindeutigkeitsaussage in Satz 2.1.8 $\mathrm{id}_{\mathbb{N}} = \psi \circ \phi$ gelten. Analog folgt $\phi \circ \psi = \mathrm{id}_{\tilde{\mathbb{N}}}$ und daher ist $\phi : \mathbb{N} \to \tilde{\mathbb{N}}$ bijektiv. $\square$

**Bemerkung 2.1.10.** Durch Anwendung des Rekursionssatzes lassen sich Addition, Multiplikation und die „kleiner“ Relation $<$ auf $\mathbb{N}$ definieren:

- Zu $m \in \mathbb{N}$ wähle $Y = \mathbb{N}$, $y = m + 1 = N(m)$ und $f = N$, dann gibt es nach Satz 2.1.8 genau eine Funktion $\varphi_m : \mathbb{N} \to \mathbb{N}$ mit $\varphi_m(1) = m + 1$ und $\varphi_m \circ N = N \circ \varphi_m$ und wir definieren dann für alle $m, n \in \mathbb{N}$:
$$m + n := \varphi_m(n) \tag{2.1.2}$$
- Zu $m \in \mathbb{N}$ wähle $Y = \mathbb{N}$, $y = m$ und $f = \varphi_m$, dann gibt es nach Satz 2.1.8 genau eine Funktion $\psi_m : \mathbb{N} \to \mathbb{N}$ mit $\psi_m(1) = m$ und $\psi_m \circ N = \varphi_m \circ \psi_m$ und wir definieren dann für alle $m, n \in \mathbb{N}$:
$$m \cdot n := \psi_m(n) \tag{2.1.3}$$
- Für $m, n \in \mathbb{N}$ schreibe $m < n$ genau dann, wenn es ein $p \in \mathbb{N}$ gibt mit $m + p = n$.

Die bekannten Rechenregeln für natürliche Zahlen benötigen jetzt einen Beweis:

**Lemma 2.1.11.** *Für alle $m, n, p \in \mathbb{N}$ gilt:*

$$(m + n) + p = m + (n + p) \tag{2.1.4}$$
$$m + n = n + m \tag{2.1.5}$$
$$(m \cdot n) \cdot p = m \cdot (n \cdot p) \tag{2.1.6}$$
$$m \cdot n = n \cdot m \tag{2.1.7}$$
$$m \cdot (n + p) = (m \cdot n) + (m \cdot p) \tag{2.1.8}$$

[2] sieht man nach dem Motto “Namen sind Schall und Rauch” von den verwendeten Namen ab – die Formulierung dessen sieht man in Lemma 2.1.9: Durch $\phi$ können die ausgezeichneten Elemente 1 und $\tilde{1}$ identifiziert werden und der Nachfolger von $\phi(n)$ mit dem Nachfolger von $n$ etc., so daß $\mathbb{N}$ und $\tilde{\mathbb{N}}$ dieselben Eigenschaften haben.

*Es gelten die* **„Kürzungsregeln“** *: Sind $m, n, p \in \mathbb{N}$, dann gilt:*

$$n + m = p + m \quad \Rightarrow \quad n = p \tag{2.1.9}$$
$$m \cdot n = p \cdot n \quad \Rightarrow \quad m = p. \tag{2.1.10}$$

*Für alle $m, n \in \mathbb{N}$ gilt entweder $m < n$ oder $m = n$ oder $n < m$. Sind $m < n$ und $n < p$, dann ist $m < p$ und aus $m < n$ folgt $m + p < n + p$ und $m \cdot p < n \cdot p$ für alle $p \in \mathbb{N}$.*

*Beweis.* Wir zeigen exemplarisch die Kürzungsregel (2.1.9): $n+m = p+m$ bedeutet nach Definition $\varphi_n(m) = \varphi_p(m)$.

- 1.Fall: $m = 1$: Dann ist $\varphi_n(1) = N(n) = \varphi_p(1) = N(p)$ und aus der Injektivität von $N$ folgt $n = p$.

- 2.Fall: $m \neq 1$. Dann [3] ist $m \in N(\mathbb{N})$ und wir (definieren $m - 1$ als) $m = N(m - 1)$. Nach Voraussetzung ist

$$\varphi_n(m) = \varphi_n(N(m-1)) = (N \circ \varphi_n)(m-1) = \varphi_p(m) = \varphi_p(N(m-1)) = (N \circ \varphi_p)(m-1),$$

  also $\varphi_n(m - 1) = \varphi_p(m - 1)$, da $N$ injektiv ist. Ist nun $m - 1 > 1$, so wiederholen wir den Schritt im 2. Fall noch $m - 2$ mal, dann sind wir im 1. Fall und bekommen $n = p$.

□

**Satz 2.1.12.** *Es sei $\emptyset \neq X \subseteq \mathbb{N}$, dann hat $X$ ein* **kleinstes Element***, dh. es gibt ein $x \in X$, so daß für alle $y \in X$ mit $y \neq x$ gilt: $x < y$.*

*Beweis.* Wir zeigen die Eindeutigkeit eines kleinsten Elements: Sind $x, x' \in X$ mit $x < n$ für alle $n \in X\backslash\{x\}$ und $x' < m$ für alle $m \in X\backslash\{x'\}$ und $x \neq x'$, dann tritt nach Lemma 2.1.11 genau einer der folgenden Fälle ein:

- $x < x'$, was im Widerspruch zur definierenden Eigenschaft $x' < m$ für alle $m \in X\backslash\{x'\}$ steht.

- $x' < x$, was im Widerspruch zur definierenden Eigenschaft $x < n$ für alle $n \in X\backslash\{x\}$ steht.

Daher ist das kleinste Element von $X$ – falls es existiert – eindeutig. Da $X \neq \emptyset$, existiert ein $m \in X$. Ist nun $X \cap \{1, ..., m\} = \{m\}$, so gilt $m < n$ für alle $n \in X\backslash\{m\}$, also ist $m$ das gesuchte kleinste Element von $X$. Gibt es ein $m' \neq m$ mit $m' \in X \cap \{1, ..., m\}$, so ist $m \notin \{1, ..., m'\} \cap X$, daher finden wir nach höchstens $m$ Schritten ein $x \in \{1, ..., m\}$ mit $\{x\} = X \cap \{1, ..., x\}$. Dann ist $x$ das kleinste Element von $X$. □

[3] Die Nachfolgeabbildung $N : \mathbb{N} \to \mathbb{N}\backslash\{1\}$ ist auch surjektiv: Betrachte dazu $X := \{1\} \cup N(\mathbb{N})$, dann ist $1 \in X$ und

$$N(X) = \{N(1)\} \cup (N \circ N)(\mathbb{N}) \subseteq \{N(1)\} \cup N(\mathbb{N}\backslash\{1\}) = N(\mathbb{N}) \subseteq X$$

und daher $X = \mathbb{N}$ nach dem Peanoaxiom (P3). Aus $1 \notin N(\mathbb{N})$ folgt nun $N(\mathbb{N}) = \mathbb{N}\backslash\{1\}$

**Satz 2.1.13** (vollständige Induktion, 2. Fassung)**.** *Es sei $n \in \mathbb{N}$ und für alle $m \geq n$ seien Aussagen $A(n, n+1, ..., m-1, m)$ gegeben. Gilt nun*

*a) $A(n)$ ist wahr.*

*b) $A(n, n+1, ..., m-1, m)$ ist wahr $\Rightarrow A(n, n+1, ..., m, m+1)$ ist wahr*

*dann ist $A(n, n+1, ..., m)$ wahr für alle $m \in \mathbb{N}$ mit $m \geq n$.*

*Beweis.* Angenommen die Folgerung ist falsch, dann ist

$$\emptyset \neq X := \{m \in \mathbb{N} : m \geq n, A(n, ..., m) \text{ ist falsch}\}.$$

Nach Satz 2.1.12 gibt es ein kleinstes Element $x \in X$. Wegen Voraussetzung a) ist dann $x > n$, $A(n, ..., x-1)$ ist wahr und $A(n, ..., x)$ falsch. Nach Voraussetzung b) gilt aber $A(n, ..., x-1)$ wahr $\Rightarrow A(n, ..., x)$ wahr; Widerspruch. □

**Beispiel 2.1.14.** Frage: Für welche $n \in \mathbb{N}$ gilt $2^n > n^2$?
Offenbar ist $2^1 = 2 > 1^2 = 1$, $2^2 = 4 = 2^2$, $2^3 = 8 < 9 = 3^2$, $2^4 = 16 = 4^2$, $2^5 = 32 > 25 = 5^2$, also die Aussage $A(n)$ gleich „$2^n > n^2$“ wahr für $n = 1, 5$ und falsch für $n = 2, 3, 4$. Wir starten nun mit dem

- Induktionsanfang: $n = 5$: $2^5 = 32 > 25 = 5^2$.
- Induktionsschluß: Sei $n \geq 5$ und $A(n)$ wahr, dann gilt:

$$2^{n+1} = 2 \cdot 2^n > 2n^2 > n^2 + 3n > n^2 + 2n + 1 = (n+1)^2$$

Daher ist $2^n > n^2$ für $n = 1$ und für alle $n \geq 5$.

## 2.2 Mächtigkeit von Mengen, Kardinalzahlen und abzählbare Mengen

**Definition 2.2.1.** *Sind $X$ und $Y$ Mengen, so heißen $X$ und $Y$* **gleichmächtig***, wenn es eine bijektive Abbildung $f : X \to Y$ gibt. Eine Menge $X$ heißt*

- **endlich***, wenn $X = \emptyset$ oder wenn es ein $n \in \mathbb{N}$ und eine bijektive Abbildung $f : X \to \{1, ..., n\}$ gibt.*
- **unendlich***, wenn $X$ nicht endlich ist.*
- **abzählbar unendlich***, wenn es eine bijektive Abbildung $f : X \to \mathbb{N}$ gibt.*
- **abzählbar***, wenn $X$ endlich oder abzählbar unendlich ist.*
- **überabzählbar***, wenn $X$ nicht abzählbar ist.*

**Bemerkung 2.2.2.** Ist $X = \{X_i : i \in I\}$ eine Menge, deren Elemente Mengen sind, so definiert „gleichmächtig“ eine Äquivalenzrelation $\sim$ auf $X$:

- Da $\mathrm{id}_{X_i} : X_i \to X_i$ bijektiv ist, gilt $X_i \sim X_i$ und daher ist $\sim$ reflexiv.
- Ist $X_i \sim X_j$, so gibt es eine bijektive Abbildung $f : X_i \to X_j$. Dann existiert auch die Umkehrabbildung $f^{-1} : X_j \to X_i$ und ist bijektiv, daher gilt $X_j \sim X_i$ und somit ist $\sim$ symmetrisch.
- Sind $X_i \sim X_j$ und $X_j \sim X_k$, so gibt es bijektive Abbildungen $f : X_i \to X_j$ und $g : X_j \to X_k$ und dann ist auch $g \circ f : X_i \to X_k$ bijektiv, also $X_i \sim X_k$ und daher $\sim$ transitiv.

Die Äquivalenzklasse $[X_i]_\sim =: \mathrm{Kard}X_i$ von $X_i$ bezüglich $\sim$ heißt die **Kardinalzahl** von $X_i$. Ist $X_i = \{x_1, ..., x_n\}$ eine endliche Menge mit $n$ Elementen, schreibt man auch $\mathrm{Kard}X_i = n$.

Bei endlichen Mengen klappt die Intuition meistens, das liegt an:

**Lemma 2.2.3.** *Es sei $n \in \mathbb{N}$ und $X, Y$ seien Mengen mit $\mathrm{Kard}X = \mathrm{Kard}Y = n$. Für eine Funktion $f : X \to Y$ sind äquivalent:*

*a) $f$ bijektiv*

*b) $f$ injektiv*

*c) $f$ surjektiv*

*Beweis.* Es genügt die Äquivalenz von b) und c) zu zeigen; schreibe $X = \{x_1, ..., x_n\}$

b) $\Rightarrow$ c) Da $f$ injektiv ist, besteht $f(X)$ aus $n$ Elementen. Mit $f(X) \subseteq Y$ folgt $f(X) = Y$, da $Y$ genau $n$ Elemente hat, also jedes Element von $Y$ auch in $f(X)$ liegt.

c) $\Rightarrow$ b) Ist $f$ surjektiv, so gilt $f(X) = Y = \{f(x_1), ..., f(x_n)\}$, also ist auch $f(x_k) \neq f(x_l)$ für $l \neq k$, dh. $f$ ist injektiv. $\square$

**Warnung 2.2.4.** Ohne die Voraussetzung, daß $X$ und $Y$ endliche Mengen sind, ist Lemma 2.2.3 falsch. Konkrete Beispiele: $\begin{aligned} f : \mathbb{N} &\to \mathbb{N} \\ n &\mapsto 2n \end{aligned}$ ist injektiv aber nicht surjektiv.
$\begin{aligned} g : \mathbb{N} &\to \mathbb{N} \\ n &\mapsto \begin{cases} \frac{n}{2} & \text{falls } n \text{ gerade ist} \\ 1 & \text{falls } n \text{ ungerade ist} \end{cases} \end{aligned}$ ist surjektiv aber nicht injektiv.

**Beispiel 2.2.5.** Jede Teilmenge $X \subseteq \mathbb{N}$ ist abzählbar. Denn wenn $X$ nicht endlich ist, dann ist

- $X \neq \emptyset$, also gibt es nach Satz 2.1.12 ein kleinstes Element $x_1 \in X$.
- $X \setminus \{x_1\} \neq \emptyset$, also gibt es nach Satz 2.1.12 ein kleinstes Element $x_2 \in X \setminus \{x_1\}$.
- $X \setminus \{x_1, x_2\} \neq \emptyset$, also gibt es nach Satz 2.1.12 ein kleinstes Element $x_3 \in X \setminus \{x_1, x_2\}$ etc.

und da $X$ unendlich ist, bricht dieses Verfahren nicht ab. Wir bekommen so eine Folge $(x_n)_{n \in \mathbb{N}}$ mit $X = \{x_n : n \in \mathbb{N}\}$ und $x_1 < x_2 < x_3 < ...$, daher ist $\begin{aligned} f : \mathbb{N} &\to X \\ n &\mapsto x_n \end{aligned}$ bijektiv.

**Lemma 2.2.6.** *Es sei $\emptyset \neq X$ eine Menge, dann sind äquivalent:*

*a) $X$ ist abzählbar.*

*b) Es gibt eine injektive Abbildung $f : X \to \mathbb{N}$.*

*c) Es gibt eine surjektive Abbildung $g : \mathbb{N} \to X$.*

*Beweis.*

a) $\Rightarrow$ b): Ist $X$ endlich, so gibt es nach Definition ein $n \in \mathbb{N}$ und eine bijektive Abbildung $\tilde{f} : X \to \{1, ..., n\}$, also ist $\begin{array}{rcl} f : X & \to & \mathbb{N} \\ x & \mapsto & \tilde{f}(x) \end{array}$ injektiv. Ist $X$ abzählbar unendlich, gibt es nach Definition eine bijektive Abbildung $f : X \to \mathbb{N}$.

b) $\Rightarrow$ a): Ist $f : X \to \mathbb{N}$ injektiv, so ist $\begin{array}{rcl} \tilde{f} : X & \to & f(X) \\ x & \mapsto & f(x) \end{array}$ bijektiv, also sind $X$ und $f(X)$ gleichmächtig. $f(X) \subseteq \mathbb{N}$ ist nach Beispiel 2.2.5 abzählbar; daher ist auch $X$ abzählbar.

a) $\Rightarrow$ c): Ist $X$ abzählbar, dann gilt im Fall

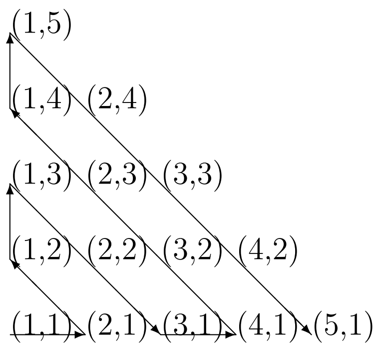
- $X$ endlich: Es gibt ein $n \in \mathbb{N}$ und eine bijektive Funktion $f : X \to \{1, ..., n\}$. Mit der surjektiven Funktion $\begin{array}{rcl} h : \mathbb{N} & \to & \{1, ..., n\} \\ k & \mapsto & \begin{cases} k & \text{für } k = 1, ..., n \\ 1 & \text{für } k \geq n \end{cases} \end{array}$ ist nach Lemma 1.3.11 auch $g = f^{-1} \circ h : \mathbb{N} \to X$ surjektiv.
- $X$ abzählbar unendlich, dann gibt es eine bijektive Funktion $f : X \to \mathbb{N}$ und $g = f^{-1}$ erfüllt die Behauptung.

c) $\Rightarrow$ a): Ist $g : \mathbb{N} \to X$ surjektiv, so gibt es nach Lemma 1.3.12 eine Funktion $f : X \to \mathbb{N}$ mit $g \circ f = \mathrm{id}_X$. Dann ist $\begin{array}{rcl} \tilde{f} : X & \to & f(X) \\ x & \mapsto & f(x) \end{array}$ bijektiv, denn aus $\tilde{f}(x) = \tilde{f}(y)$ folgt $x = g(f(x)) = g(f(y)) = y$. Nach Beispiel 2.2.5 ist $f(X) \subseteq \mathbb{N}$ abzählbar, also auch die gleichmächtige Menge $X$.

□

**Satz 2.2.7.** *Sind $X$ und $Y$ abzählbare Mengen, so ist auch $X \times Y$ abzählbar.*

*Beweis.* Ohne Einschränkung können wir $X \neq \emptyset$ und $Y \neq \emptyset$ annehmen. Nach Lemma 2.2.6 gibt es injektive Abbildungen $f_X : X \to \mathbb{N}$ und $f_Y : Y \to \mathbb{N}$, dann ist auch $\begin{array}{rcl} f : X \times Y & \to & \mathbb{N} \times \mathbb{N} \\ (x, y) & \mapsto & (f_X(x), f_Y(y)) \end{array}$ injektiv – denn nach Definition der Gleichheit in cartesischen Produkten gilt zu $(x_1, y_1), (x_2, y_2) \in \mathbb{N} \times \mathbb{N}$ die Gleichheit $(f_X(x_1), f_Y(y_1)) = (f_X(x_2), f_Y(y_2))$ genau dann wenn $f_X(x_1) = f_X(x_2)$ und $f_Y(y_1) = f_Y(y_2)$ ist. Wegen Injektivität von $f_X$ und $f_Y$ folgt daraus $x_1 = x_2$ und $y_1 = y_2$, also $(x_1, y_1) = (x_2, y_2)$. Weil die Komposition injektiver Funktionen nach Lemma 1.3.11 wieder injektiv ist, reicht es nach Lemma 2.2.6 zu zeigen, daß $\mathbb{N} \times \mathbb{N}$ abzählbar ist. Hier verwenden wir das Cantorsche Diagonalverfahren




das uns eine surjektive Funktion $g : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ gibt: Zuerst durchläuft man alle Tupel $(n, m) \in \mathbb{N}^2$ mit $n + m = 2$, dann die mit $n + m = 3$, dann die mit $n + m = 4$ usw. ... Da es immer gerade $N - 1$ Tupel $(n, m) \in \mathbb{N}^2$ mit $n + m = N$ gibt, erreicht man auf diese Weise irgendwann jedes Tupel $(n, m) \in \mathbb{N}^2$. [4] □

**Korollar 2.2.8.** *Die Vereinigung von abzählbar vielen abzählbaren Mengen ist abzählbar.*

*Beweis.* Es sei $I$ eine abzählbare Menge und für alle $i \in I$ sei $X_i$ abzählbar. Dann gibt es nach Lemma 2.2.6 eine surjektive Abbildung $f : \mathbb{N} \to I$ und für alle $i \in I$ gibt es surjektive Abbildungen $g_i : \mathbb{N} \to X_i$. Dann ist

$$\begin{aligned} F : \mathbb{N} \times \mathbb{N} &\to \bigcup_{i \in I} X_i \\ (m, n) &\mapsto g_{f(m)}(n) \end{aligned}$$

ebenfalls surjektiv, denn zu jedem $x \in \bigcup_{i \in I} X_i$ gibt es ein $j \in I$ mit $x \in X_j$. Da $f$ und $g_j$ surjektiv sind, gibt es $m \in \mathbb{N}$ mit $j = f(m)$ und ein $n \in \mathbb{N}$ mit $g_j(n) = x$, also ist $x = g_{f(m)}(n) = F(m, n)$. Nach Satz 2.2.7 ist $\mathbb{N} \times \mathbb{N}$ abzählbar, daher gibt es nach Lemma 2.2.6 eine surjektive Abbildung $h : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$. Nach Lemma 1.3.11 ist $F \circ h : \mathbb{N} \to \bigcup_{i \in I} X_i$ surjektiv, also $\bigcup_{i \in I} X_i$ abzählbar. □

**Lemma 2.2.9.** $\mathcal{P}(\mathbb{N})$ *ist überabzählbar.*

*Beweis.* Zu $g : \mathbb{N} \to \mathcal{P}(\mathbb{N})$ betrachte $Y := \{n \in \mathbb{N} : n \notin g(n)\} \subseteq \mathbb{N}$. Für $n \in \mathbb{N}$ mit $Y = g(n)$ ist die Bedingung $n \in Y$ mit $n \notin Y$ äquivalent. Aus diesem Widerspruch folgt $Y \notin g(\mathbb{N})$, also ist $g$ nicht surjektiv. Damit ist $\mathcal{P}(\mathbb{N})$ überabzählbar, denn andernfalls müßte es nach Lemma 2.2.6 eine surjektive Abbildung $g : \mathbb{N} \to \mathcal{P}(\mathbb{N})$ geben. □

---

[4] Man könnte hier eine ganz explizite Definition von $g$ angeben – aber damit die Surjektivität zu zeigen würde die Sache nicht klarer machen.

# Kapitel 3

# Algebraische Grundstrukturen

## 3.1 Gruppen

**Definition 3.1.1.** *Eine* **Gruppe** $(G, *)$ *besteht aus einer Menge* $G$ *zusammen mit einer Abbildung*

$$\begin{array}{rcl} * : G \times G & \to & G \\ (g, h) & \mapsto & g * h \end{array}$$

*die folgende Eigenschaften besitzt:*

*a)* **Assoziativität:** *für alle* $g, h, j \in G$ *gilt*

$$(g * h) * j = g * (h * j)$$

*– wir schreiben dann* $g * h * j$.

*b) Es gibt ein* **linksneutrales Element** $e \in G$ *mit* $e * g = g$ *für alle* $g \in G$.

*c) Für alle* $g \in G$ *gibt es ein* **linksinverses Element** $g' \in G$ *mit* $g' * g = e$.

*Eine Gruppe* $(G, *)$ *heißt* **kommutativ**, *wenn* $g * h = h * g$ *für alle* $g, h \in G$ *gilt.*

**Lemma 3.1.2.** *Es sei* $(G, *)$ *eine Gruppe, dann gilt:*

*a) Das linksneutrale Element* $e \in G$ *ist eindeutig bestimmt und es gilt auch* $g * e = g$ *für alle* $g \in G$*; deshalb heißt* $e$ **neutrales Element** *von* $(G, *)$.

*b) Das linksinverse Element* $g' \in G$ *zu* $g \in G$ *ist eindeutig bestimmt und es gilt ferner* $g * g' = e$*; wir schreiben dann* $g^{-1}$ *für das* **inverse Element** *zu* $g \in G$.

*c) In* $(G, *)$ *gelten die Rechenregeln: Für* $g, h \in G$ *gilt:*

$$g^{-1} * g = g * g^{-1} = e \quad (3.1.1)$$

$$(g^{-1})^{-1} = g \quad (3.1.2)$$

$$(g * h)^{-1} = h^{-1} * g^{-1} \quad (3.1.3)$$

d) *In* $(G, *)$ *gelten die* **Kürzungsregeln:** *Für* $g, h, j \in G$ *gilt:*

$$g * h = g * j \quad \Rightarrow \quad h = j \tag{3.1.4}$$
$$g * h = j * h \quad \Rightarrow \quad g = j \tag{3.1.5}$$

*Beweis.* Es sei $e * g = g$ für alle $g \in G$ und es seien $g', g'' \in G$ mit $g' * g = e$ und $g'' * g' = e$. Dann folgt:

$$g * g' = e * (g * g') = (g'' * g') * (g * g') = g'' * (g' * g) * g' = g'' * (e * g') = g'' * g' = e \tag{3.1.6}$$

und damit

$$g * e = g * (g' * g) = (g * g') * g = e * g = g,$$

dh. jedes linksneutrale Element $e \in G$ ist ein neutrales Element in $G$. Sind nun $e$ und $\tilde{e}$ neutrale Elemente in $G$, so gilt $e * \tilde{e} = e$ (da $\tilde{e}$ neutral ist) und $e * \tilde{e} = \tilde{e}$ (da $e$ neutral ist), also $e = \tilde{e}$, das zeigt a). Damit zeigt (3.1.6), daß ein linksinverses Element $g'$ von $g$ auch ein inverses Element zu $g$ ist. Da für jedes inverse Element $\tilde{g}$ von $g$ gilt

$$\tilde{g} = \tilde{g} * e = \tilde{g} * (g * g') = (\tilde{g} * g) * g' = e * g' = g'$$

folgt die Eindeutigkeit des inversen Elements $g^{-1}$ von $g$. Aus $g * g^{-1} = e$ folgt, daß $g$ inverses Element zu $g^{-1}$ ist, daher $(g^{-1})^{-1} = g$ nach a) und b). Ferner gilt für $g, h \in G$:

$$(h^{-1} * g^{-1}) * (g * h) = h^{-1} * (g^{-1} * g) * h = h^{-1} * e * h = h^{-1} * h = e$$

und damit $(g * h)^{-1} = h^{-1} * g^{-1}$. Die Kürzungsregel (3.1.4) folgt durch Verknüpfung mit $g^{-1}$ von links: $h = g^{-1} * (g * h) = g^{-1} * (g * j) = j$ und (3.1.5) folgt durch Verknüpfung mit $h^{-1}$ von rechts: $g = (g * h) * h^{-1} = (j * h) * h^{-1} = j$. □

**Beispiel 3.1.3.**

1. Ist $\emptyset \neq X$ eine Menge, so ist

   $$\mathcal{S}(X) := \{f : X \to X : f \text{ bijektiv }\}$$

   mit der Komposition von Funktionen als Verknüpfung eine Gruppe:

   - Nach Lemma 1.3.6 ist $\circ$ assoziativ.
   - $\mathrm{id}_X$ ist das neutrale Element in $\mathcal{S}(X)$.
   - Die Umkehrfunktion $f^{-1}$ zu $f \in \mathcal{S}(X)$ ist das inverse Element zu $f$ in $(\mathcal{S}(X), \circ)$.

2. Ist $X$ eine Menge mit $\mathrm{Kard} X \geq 2$, so ist $\mathrm{Abb}(X, X) := \{f : X \to X\}$ zusammen mit $\circ$ keine Gruppe, denn wenn $x_0 \in X$ ist, so ist $\begin{array}{rccc} f_{x_0} : & X & \to & X \\ & x & \mapsto & x_0 \end{array}$ nicht bijektiv, also gibt es kein inverses Element zu $f_{x_0}$.

3. $(\mathbb{N}, +)$ ist keine Gruppe, denn für alle $m, n \in \mathbb{N}$ ist $m + n > n$, daher gibt es in $(\mathbb{N}, +)$ kein neutrales Element.

4. $(\mathbb{N}, \cdot)$ ist keine Gruppe, denn $1 \in \mathbb{N}$ ist zwar ein neutrales Element, aber wegen $m \cdot n > n$ für alle $m, n \in \mathbb{N}$ mit $m \geq 2$ genügt es etwa $m = 2 > 1$ zu betrachten. Wegen $2 \cdot n > n \geq 1$ hat 2 in $(\mathbb{N}, \cdot)$ kein inverses Element.

5. Für $X = \{1, ..., n\}$ ist $\mathcal{S}_n := \mathcal{S}(X)$ eine Gruppe. Ein Element $\pi \in \mathcal{S}_n$ heißt **Permutation**; man schreibt es üblicherweise in der Form $\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}$. In der ersten Zeile steht das Element des Definitionsbereichs $\{1, ..., n\}$ und in der zweiten Zeile darunter der Funktionswert. Etwa für $n = 3$ ist

$$\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{array}{rcl} \pi : \{1,2,3\} & \to & \{1,2,3\} \\ 1 & \mapsto & 2 \\ 2 & \mapsto & 3 \\ 3 & \mapsto & 1 \end{array}$$

Für $n \geq 3$ erhalten wir mit $(\mathcal{S}_n, \circ)$ Beispiele für Gruppen, die nicht kommutativ sind, etwa für $\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ und $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ ist

$$\pi \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ \pi(\sigma(1)) & \pi(\sigma(2)) & \pi(\sigma(3)) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \neq \sigma \circ \pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

**Lemma 3.1.4.** *Es sei $G = \{g_1, ..., g_n\}$ eine endliche Menge und $* : G \times G \to G$ assoziativ. Dann sind äquivalent:*

*a)* *$(G, *)$ ist eine Gruppe*

*b)* *In $(G, *)$ gelten die Kürzungsregeln*

$$\begin{array}{lcl} g * h = g * j & \Rightarrow & h = j \\ g * h = j * h & \Rightarrow & g = j \end{array}$$

*c)* *Für jedes $g \in G$ sind*

$$\begin{array}{rcl} l_g : G & \to & G \\ h & \mapsto & g * h \end{array} \quad \text{und} \quad \begin{array}{rcl} r_g : G & \to & G \\ h & \mapsto & h * g \end{array}$$

*bijektiv.*

*d)* *In der* **Verknüpfungstafel**

| $*$ | $g_1$ | $g_2$ | $\cdots$ | $g_n$ |
|---|---|---|---|---|
| $g_1$ | $g_1 * g_1$ | $g_1 * g_2$ | $\cdots$ | $g_1 * g_n$ |
| $g_2$ | $g_2 * g_1$ | $g_2 * g_2$ | $\cdots$ | $g_2 * g_n$ |
| $\vdots$ | $\vdots$ | | | |
| $g_n$ | $g_n * g_1$ | $g_n * g_2$ | $\cdots$ | $g_n * g_n$ |

*zu $(G, *)$ kommt in jeder Zeile und in jeder Spalte jedes Element von $G$ genau einmal vor.*

*Beweis.*

$a) \Rightarrow b)$ nach Lemma 3.1.2.

$b) \Rightarrow c)$ Gelten die Kürzungsregeln, so sind für jedes $g \in G$ die Funktionen $l_g$ und $r_g$ injektiv. Da $G$ endlich ist, sind $l_g$ und $r_g$ nach Lemma 2.2.3 bijektiv.

$c) \Rightarrow a)$ Sind $l_g$ und $r_g$ für alle $g \in G$ bijektiv, so hat jede der Gleichungen

$$g * x = h \qquad\qquad y * g = h$$

eine eindeutige Lösung. Insbesondere gibt es (bei Wahl von $g = h$) ein $e = e(g) \in G$ mit $e * g = g$. Ferner gibt es für alle $j \in G$ ein $k \in G$ mit $g * k = j$, also ist $e * j = e * (g * k) = (e * g) * k = g * k = j$, daher ist $e$ linksneutrales Element in $(G, *)$ und die Lösung von $y * g = e$ gibt ein Linksinverses zu $g$, dh. alle Gruppenaxiome für $(G, *)$ gelten.

$c) \Leftrightarrow d)$ Da in der $j$-ten Spalte der Verknüpfungstafel die Bilder von $r_{g_j}$ und in der $i$-ten Zeile der Verknüpfungstafel die Bilder von $l_{g_i}$ stehen, folgt die Äquivalenz aus Lemma 2.2.3. $\square$

**Beispiel 3.1.5.** Lemma 3.1.4 erlaubt uns die möglichen Gruppen auf einer Menge mit endlich vielen Elementen zu bestimmen.

- Für eine zweielementige Menge gibt es nur eine Möglichkeit diese zu einer Gruppe zu machen: Da eine Gruppe ein neutrales Element besitzen muß – dieses wollen wir jetzt mit $e$ bezeichnen – schreiben wir $G = \{e, a\}$. Nach den Rechenregeln muß ferner $e * a = a * e = a$ und $e * e = e$ sein, also können wir die Verknüpfungstafel bereits so weit ausfüllen:

  | $*$ | $e$ | $a$ |
  |---|---|---|
  | $e$ | $e$ | $a$ |
  | $a$ | $a$ | ? |

  und nach Lemma 3.1.4 muß bei ? ein $e$ eingefügt werden um $(G, *)$ zu einer Gruppe zu machen. Nachprüfen ergibt, daß $*$ eine assoziative Verknüpfung ist, damit wird $(G, *)$ eine Gruppe.

- Auch auf einer Menge $G = \{e, a, b\}$ mit drei Elementen gibt es genau eine Gruppenstruktur. Wie eben können wir die Verknüpfungstafel so weit ausfüllen:

  | $*$ | $e$ | $a$ | $b$ |
  |---|---|---|---|
  | $e$ | $e$ | $a$ | $b$ |
  | $a$ | $a$ | ? | |
  | $b$ | $b$ | ?? | |

  Bei ? kann also laut Lemma 3.1.4 kein $a$ stehen und wenn man bei ? ein $e$ einträgt, müßte – bei Betrachten der 2. Spalte – statt ?? ein $b$ eingefügt werden, was aber –

bei Betrachten der 3. Zeile nicht geht. Bei ? muß also ein $b$ eingefügt werden und dann ergänzt sich die Verknüpfungstafel wie von selbst:

| $*$ | $e$ | $a$ | $b$ |
|---|---|---|---|
| $e$ | $e$ | $a$ | $b$ |
| $a$ | $a$ | $b$ | $e$ |
| $b$ | $b$ | $e$ | $a$ |

und man braucht nur noch nachzurechnen, daß $*$ wirklich assoziativ ist.

**Satz 3.1.6.** *Auf* $\mathbb{N} \times \mathbb{N}$ *wird für* $k, l, m, n \in \mathbb{N}$ *durch*

$$(k, l) \sim (m, n) :\Leftrightarrow k + n = l + m$$

*eine Äquivalenzrelation* $\sim$ *definiert. Auf*

$$\mathbb{Z} := (\mathbb{N} \times \mathbb{N})/_\sim$$

*wird durch*

$$[(k, l)]_\sim + [(m, n)]_\sim := [(k + m, l + n)]_\sim \tag{3.1.7}$$

*eine Abbildung* $+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ *definiert, so daß* $(\mathbb{Z}, +)$ *eine kommutative Gruppe ist.*

*Beweis.* Durch (3.1.7) wird eine Abbildung $+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ definiert, denn für $(k_1, l_1)$, $(k_2, l_2), (m_1, n_1), (m_2, n_2) \in \mathbb{N}\times\mathbb{N}$ mit $[(k_1, l_1)]_\sim = [(k_2, l_2)]_\sim$ und $[(m_1, n_1)]_\sim = [(m_2, n_2)]_\sim$, also $k_1 + l_2 = k_2 + l_1$ und $m_1 + n_2 = m_2 + n_1$ gilt

$$[(k_1{+}l_1)]_\sim{+}[(m_1, n_1)]_\sim = [(k_1{+}m_1, l_1{+}n_1)]_\sim = [(k_2{+}l_2)]_\sim{+}[(m_2, n_2)]_\sim = [(k_2{+}m_2, l_2{+}n_2)]_\sim,$$

da $k_1 + m_1 + l_2 + n_2 = l_1 + n_1 + k_2 + m_2$.

- Assoziativität: Ist $(k_1, l_1), (k_2, l_2), (k_3, l_3) \in \mathbb{N} \times \mathbb{N}$, so gilt:

$$\begin{aligned} [(k_1, l_1)]_\sim + \Big([(k_2, l_2)]_\sim + [(k_3, l_3)]_\sim\Big) &= [(k_1, l_1)]_\sim + [(k_2 + k_3, l_2 + l_3)]_\sim \\ = \ [(k_1 + k_2 + k_3, l_1 + l_2 + l_3)]_\sim &= [(k_1 + k_2, l_1 + l_2)]_\sim + [(k_3, l_3)]_\sim \\ = \ \Big([(k_1, l_1)]_\sim + [(k_2, l_2)]_\sim\Big) + [(k_3, l_3)]_\sim & \end{aligned}$$

- Kommutativität: Ist $(k_1, l_1), (k_2, l_2) \in \mathbb{N} \times \mathbb{N}$, so gilt:

$$[(k_1, l_1)]_\sim{+}[(k_2, l_2)]_\sim = [(k_1{+}k_2, l_1{+}l_2)]_\sim = [(k_2{+}k_1, l_2{+}l_1)]_\sim = [(k_2, l_2)]_\sim{+}[(k_1, l_1)]_\sim$$

- $[(1, 1)]_\sim$ ist neutrales Element, denn zu $(k, l) \in \mathbb{N} \times \mathbb{N}$ ist $[(k, l)]_\sim + [(1, 1)]_\sim = [(k + 1, l + 1)]_\sim = [(k, l)]_\sim$, wobei die letzte Gleichheit aus $(k + 1) + l = (l + 1) + k$ folgt.

- Zu $[(m, n)]_\sim \in \mathbb{Z}$ ist $[(n, m)]_\sim$ das inverse Element in $(\mathbb{Z}, +)$, denn

$$[(m, n)]_\sim + [n, m)]_\sim = [(m + n, n + m)]_\sim = [(1, 1)]_\sim.$$

$\square$

**Bemerkung 3.1.7.** Die Abbildung $\begin{array}{rcl} i : \mathbb{N} & \to & \mathbb{Z} \\ m & \mapsto & [(m+1,1)]_\sim \end{array}$ ist injektiv und es gilt $i(m+n) = i(m) + i(n)$ für alle $m, n \in \mathbb{N}$. Wir können also $\mathbb{N}$ mit $i(\mathbb{N})$ identifizieren (und auch das Rechnen mit $+$ retten) und damit $\mathbb{N}$ als Teilmenge von $\mathbb{Z}$ auffassen. In der üblichen Schreibweise

$$\begin{array}{rcl} m & \equiv & [(m+1,1)]_\sim \\ 0 & := & [(1,1)]_\sim \\ -m & := & [(1,m+1)]_\sim \end{array}$$

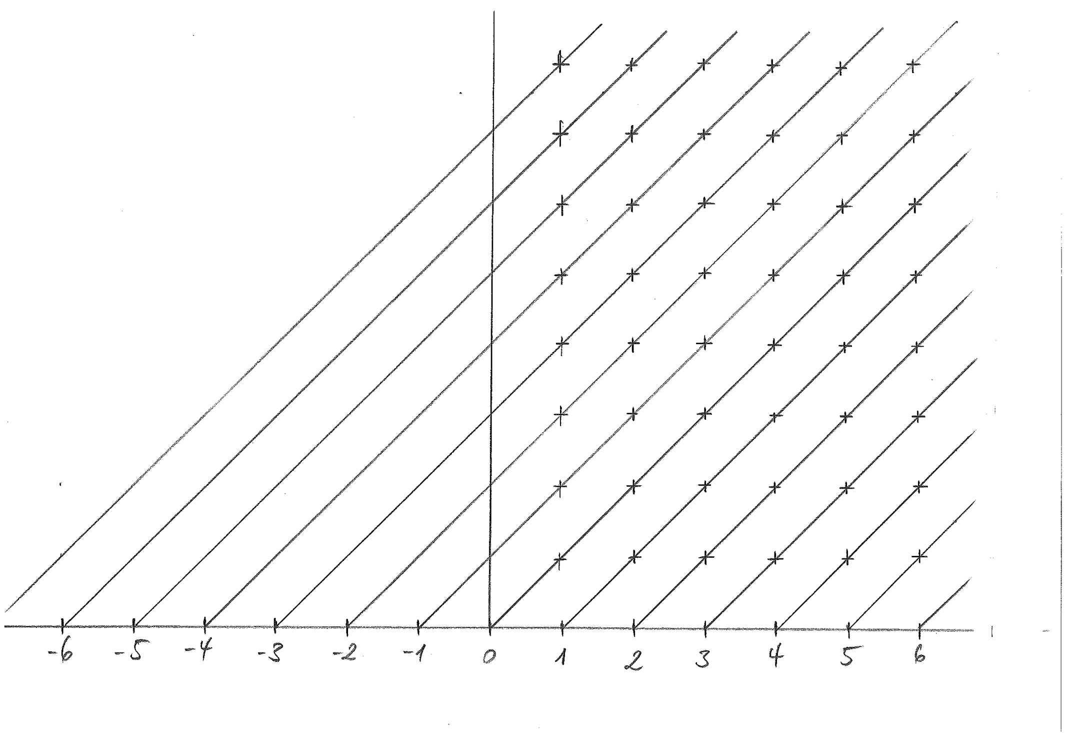
ist $-m$ inverses Element von $m \equiv i(m)$ in $(\mathbb{Z}, +)$ und es gilt

$$\mathbb{Z} = \{0\} \cup \mathbb{N} \cup \{-m : m \in \mathbb{N}\} \tag{3.1.8}$$

*Beweis.* Sind $m, n \in \mathbb{N}$ mit $[(m+1,1)]_\sim = [(n+1,1)]_\sim$, so gilt $(m+1)+1 = 1+(n+1)$, also $m = n$ nach den Kürzungsregeln in Lemma 2.1.11, und damit ist $i$ injektiv. Durch Einsetzen folgt $i(m)+i(n) = i(m+n)$ und nach Satz 3.1.6 sind $-m = [(1,m+1)]_\sim$ und $0 = [(1,1)]_\sim$ wohldefinierte Elemente von $\mathbb{Z}$ mit $i(m) + (-m) = 0$. Ist nun $[(k,l)]_\sim \in \mathbb{N} \times \mathbb{N}/_\sim$, so sind nach Lemma 2.1.11 nur folgende Fälle möglich:

- $k = l$: Dann folgt $k+1 = l+1$, also $[(k,l)]_\sim = [(1,1)]_\sim = 0$.
- $k < l$: Nach Definition von $<$ gibt es dann ein $m \in \mathbb{N}$ mit $k+m = l$, also $k+(m+1) = l+1$ oder $[(k,l)]_\sim = [(1,m+1)]_\sim \equiv -m$.
- $l < k$: Dann gibt es $m \in \mathbb{N}$ mit $l+m = k$, also $k+1 = l+(m+1)$ oder $[(k,l)]_\sim = [(m+1,1)]_\sim = m$

$\square$



**Bemerkung 3.1.8.** Die schrägen Halbgeraden im obigen Bild zeigen, welche Tupel $(k, l) \in \mathbb{N} \times \mathbb{N}$ zur selben Äquivalenzklasse gehören – hier haben sie also die gleiche Differenz – und mit welcher ganzen Zahl diese identifiziert wird.

**Definition 3.1.9.** *Sind $(G, *)$ und $(H, \odot)$ Gruppen, dann heißt $\varphi : G \to H$ ein* **Gruppenhomomorphismus**, *wenn für alle $g_1, g_2 \in G$ gilt:*

$$\varphi(g_1 * g_2) = \varphi(g_1) \odot \varphi(g_2). \tag{3.1.9}$$

*Eine nichtleere Menge $G' \subseteq G$ heißt* **Untergruppe** *von $G$, wenn für $g_1, g_2 \in G'$ auch $g_1^{-1} \in G'$ und $g_1 * g_2 \in G'$ folgt.*

**Lemma 3.1.10.** *Es seien $(G, *)$ und $(H, \odot)$ Gruppen und $\varphi : G \to H$ ein Gruppenhomomorphismus. Dann gilt:*

*a)* $\varphi(e_G) = e_H$

*b)* $\varphi(g^{-1}) = (\varphi(g))^{-1}$ *für alle* $g \in G$

*c)* $\varphi^{-1}(\{e_H\})$ *ist Untergruppe von* $G$.

*d)* $\varphi(G)$ *ist Untergruppe von* $H$.

*Beweis.*

a) Nach Definition der neutralen Elemente $e_G$ von $G$ und $e_H$ von $H$ und der Eigenschaft (3.1.9) eines Gruppenhomomorphismuses gilt:

$$e_H \odot \varphi(e_G) = \varphi(e_G) = \varphi(e_G * e_G) = \varphi(e_G) \odot \varphi(e_G),$$

woraus $e_H = \varphi(e_G)$ mit den Kürzungsregeln in Lemma 3.1.2 folgt.

b) Damit folgt

$$e_H = \varphi(e_G) = \varphi(g^{-1} * g) = \varphi(g^{-1}) \odot \varphi(g),$$

wonach $\varphi(g^{-1})$ linksinverses Element zu $\varphi(g)$, also nach Lemma 3.1.2 auch inverses Element $(\varphi(g))^{-1} = \varphi(g^{-1})$ ist.

c) Nach Teil a) ist $e_G \in \varphi^{-1}(\{e_H\}) \neq \emptyset$. Sind $g_1, g_2 \in \varphi^{-1}(\{e_H\})$, also $\varphi(g_1) = \varphi(g_2) = e_H$, dann gilt:

$$\varphi(g_1 * g_2) = \varphi(g_1) \odot \varphi(g_2) = e_H \odot e_H = e_H,$$

also $g_1 * g_2 \in \varphi^{-1}(\{e_H\})$ und

$$\varphi(g_1^{-1}) = (\varphi(g_1))^{-1} = e_H^{-1} = e_H,$$

also $g_1^{-1} \in \varphi^{-1}(\{e_H\})$.

d) $e_H = \varphi(e_G) \in \varphi(G) \neq \emptyset$ und für $h_1, h_2 \in \varphi(G)$ gibt es $g_1, g_2 \in G$ mit $h_1 = \varphi(g_1)$ und $h_2 = \varphi(g_2)$ und daher ist

$$\begin{aligned} h_1 \odot h_2 &= \varphi(g_1) \odot \varphi(g_2) = \varphi(g_1 * g_2) \in \varphi(G) \\ h_1^{-1} &= (\varphi(g_1))^{-1} = \varphi(g_1^{-1}) \in \varphi(G) \end{aligned}$$

$\square$

**Beispiel 3.1.11.** In der Gruppe

$$\begin{aligned} \mathcal{S}_3 \;=\; &\left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \right. \\ &\left. \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\} \end{aligned}$$

der Permutationen von $1, 2, 3$ ist $U = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}$ eine Untergruppe. Denn wegen

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$$

ist $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ und $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}^{-1}$ und damit sind schon alle Bedingungen an eine Untergruppe für $U$ nachgewiesen.

**Beispiel* 3.1.12.**

1. Ist $m \in \mathbb{N}$ so ist $\begin{array}{rcl} \varphi : \mathbb{Z} & \to & \mathbb{Z} \\ k & \mapsto & mk \end{array}$ ein Gruppenhomomorphismus – denn nach dem Distributivgesetz gilt $\varphi(k + l) = m(k + l) = mk + ml = \varphi(k) + \varphi(l)$ – und daher ist nach Lemma 3.1.10 $\varphi(\mathbb{Z}) = m\mathbb{Z} = \{mk : k \in \mathbb{Z}\}$ eine Untergruppe.

2. Aus den Rechenregeln der Exponentialfunktion sieht man, daß $\begin{array}{rcl} \exp : \mathbb{R} & \to & ]0, \infty[ \\ x & \mapsto & e^x \end{array}$ ein Gruppenhomomorphismus von $(\mathbb{R}, +)$ nach $(]0, \infty[, \cdot)$ ist.

**Definition 3.1.13.** *Zu $\pi \in \mathcal{S}_n$ ist*

$$\operatorname{sign}(\pi) := \prod_{\substack{1 \le i,j \le n \\ i<j}} \frac{\pi(j) - \pi(i)}{j - i} \tag{3.1.10}$$

*das* **Signum** *von $\pi$. Ein $\tau \in \mathcal{S}_n$ heißt* **Transposition***, wenn es $k, l \in \{1, ..., n\}$ mit $k \neq l$ gibt, so daß $\tau(k) = l$, $\tau(l) = k$ und $\tau(j) = j$ für alle $j \in \{1, ..., n\}$, $j \neq k$, $j \neq l$.*

**Bemerkung 3.1.14.** Für eine Transposition $\tau \in \mathcal{S}_n$ gilt $\tau = \tau^{-1}$.

**Lemma 3.1.15.** *Für $n \geq 2$ gibt es zu jedem $\pi \in \mathcal{S}_n$ Transpositionen $\tau_1, ..., \tau_r$ mit $\pi = \tau_1 \circ ... \circ \tau_r$.*

*Beweis.*

- Ist $\pi = \mathrm{id}$ und $\tau = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ 2 & 1 & 3 & \cdots & n \end{pmatrix}$, dann ist $\mathrm{id} = \tau \circ \tau^{-1} = \tau \circ \tau$.
- Ist $\pi \neq \mathrm{id}$, dann gibt es ein eindeutig bestimmtes $j_1 \in \{1, ..., n\}$ mit $\pi(1) = 1, ..., \pi(j_1 - 1) = j_1 - 1$ und $\pi(j_1) > j_1$. Wähle die Transposition $\tau_1$, die $j_1$ mit $\pi(j_1)$ vertauscht, dann ist für $\pi_1 := \tau_1 \circ \pi$

  $$\pi_1(1) = 1, ..., \pi_1(j_1) = j_1$$

  und es gibt zwei Möglichkeiten:

  a) $\pi_1 = \mathrm{id}$, dann ist $\mathrm{id} = \tau_1 \circ \pi$, dh. $\pi = \tau_1^{-1} = \tau_1$.

  b) Es gibt $j_2 > j_1$ mit

  $$\pi_1(1) = 1, ..., \pi_1(j_2 - 1) = j_2 - 1, \pi_1(j_2) > j_2.$$

  Wähle nun die Transposition $\tau_2$, die $\pi_1(j_2)$ und $j_2$ vertauscht. Dann läßt sich der Schritt wiederholen bis man (nach höchstens $n$ Schritten) zum Fall a) kommt.

  Damit kommt man zu Transpositionen $\tau_1, ..., \tau_r$ mit $\mathrm{id} = \tau_r \circ ... \circ \tau_1 \circ \pi$, dh. $\pi = (\tau_r \circ ... \circ \tau_1)^{-1} = \tau_1^{-1} \circ ... \circ \tau_r^{-1} = \tau_1 \circ ... \circ \tau_r$

□

**Lemma 3.1.16.** *Für alle $\sigma, \pi \in \mathcal{S}_n$ gilt:*

*a)* $\mathrm{sign}(\pi) \in \{-1, 1\}$

*b)* $\mathrm{sign}(\sigma \circ \pi) = \mathrm{sign}(\sigma) \cdot \mathrm{sign}(\pi)$

*c)* $\mathrm{sign}(id) = 1$

*d)* $\mathrm{sign}(\pi^{-1}) = \mathrm{sign}(\pi)^{-1} = \mathrm{sign}(\pi)$

*e)* $\mathrm{sign}(\tau) = -1$ *für jede Transposition $\tau \in \mathcal{S}_n$.*

*Beweis.* Für $i, j \in \{1, ..., n\}$, $i < j$ und $\pi \in \mathcal{S}_n$ ist $\pi(j) \neq \pi(i)$, deshalb bekommen wir durch Erweitern

$$\begin{aligned} \mathrm{sign}(\sigma \circ \pi) &= \prod_{\substack{1 \leq i,j \leq n \\ i<j}} \frac{\sigma(\pi(j)) - \sigma(\pi(i))}{j - i} = \prod_{\substack{1 \leq i,j \leq n \\ i<j}} \frac{\sigma(\pi(j)) - \sigma(\pi(i))}{\pi(j) - \pi(i)} \cdot \frac{\pi(j) - \pi(i)}{j - i} \\ &= \mathrm{sign}(\pi) \cdot \prod_{\substack{1 \leq i,j \leq n \\ i<j}} \frac{\sigma(\pi(j)) - \sigma(\pi(i))}{\pi(j) - \pi(i)} \end{aligned}$$

Forme den zweiten Term weiter um:

$$\begin{aligned}\prod_{\substack{1\le i,j\le n\\ i<j}}\frac{\sigma(\pi(j))-\sigma(\pi(i))}{\pi(j)-\pi(i)} &= \prod_{\substack{1\le i,j\le n\\ i<j\\ \pi(i)<\pi(j)}}\frac{\sigma(\pi(j))-\sigma(\pi(i))}{\pi(j)-\pi(i)}\cdot\prod_{\substack{1\le i,j\le n\\ i<j\\ \pi(i)>\pi(j)}}\frac{\sigma(\pi(j))-\sigma(\pi(i))}{\pi(j)-\pi(i)}\\ &= \prod_{\substack{1\le i,j\le n\\ i<j\\ \pi(i)<\pi(j)}}\frac{\sigma(\pi(j))-\sigma(\pi(i))}{\pi(j)-\pi(i)}\cdot\prod_{\substack{1\le i,j\le n\\ i>j\\ \pi(i)<\pi(j)}}\frac{\sigma(\pi(j))-\sigma(\pi(i))}{\pi(j)-\pi(i)}\end{aligned}$$

denn zu $(i,j)\in\{1,...,n\}^2$ mit $i<j$ und $\pi(i)>\pi(j)$ kommen wir zu $(j,i)\in\{1,...,n\}^2$ mit $j>i$ und $\pi(j)<\pi(i)$ und anschließendes Vertauschen von $i$ und $j$ zeigt die Gleichheit der Indexmengen der beiden letzten Produkte. Durch das Zusammenfassen der beiden Indexmengen folgt

$$\operatorname{sign}(\sigma\circ\pi)=\operatorname{sign}(\pi)\cdot\prod_{\substack{1\le i,j\le n\\ \pi(i)<\pi(j)}}\frac{\sigma(\pi(j))-\sigma(\pi(i))}{\pi(j)-\pi(i)}=\operatorname{sign}(\pi)\cdot\operatorname{sign}(\sigma),$$

denn da $\pi$ bijektiv ist, kommen im letzten Produkt alle Faktoren $\dfrac{\sigma(j)-\sigma(i)}{j-i}$, $i<j$ nur in einer anderen Reihenfolge vor. Das zeigt b) und Einsetzen in die Definition ergibt $\operatorname{sign}(\mathrm{id})=1$.

Ist $\tau\in\mathcal{S}_n$ die Transposition, die $k$ und $l$, $k<l$ vertauscht, dann ist $\operatorname{sign}(\tau)=\frac{\tau(l)-\tau(k)}{l-k}=-1$. Nach Lemma 3.1.15 gibt es für jedes $\pi\in\mathcal{S}_n$ Transpositionen $\tau_1,...,\tau_r$ mit $\pi=\tau_1\circ...\circ\tau_r$, also ist $\operatorname{sign}(\pi)=(-1)^r\in\{-1,1\}$, wie es durch wiederholtes Anwenden von b) folgt.

Damit ist $1=\operatorname{sign}(\mathrm{id})=\operatorname{sign}(\pi\circ\pi^{-1})=\operatorname{sign}(\pi)\cdot\operatorname{sign}(\pi^{-1})=\operatorname{sign}(\pi^{-1})\cdot\operatorname{sign}(\pi)$, also $\operatorname{sign}(\pi^{-1})$ inverses Element von $\operatorname{sign}(\pi)$. Da $g=g^{-1}$ für alle $g\in\{-1,1\}$ gilt, ist auch d) gezeigt. □

**Beispiel 3.1.17.** $\begin{array}{rcl}\operatorname{sign}:\mathcal{S}_n & \to & \{-1,1\}\\ \pi & \mapsto & \operatorname{sign}(\pi)\end{array}$ ist ein Gruppenhomomorphismus.

## 3.2 Ringe und Körper

**Definition 3.2.1.** *Eine Menge $R$ versehen mit zwei Verknüpfungen $+:R\times R\to R$ und $\cdot:R\times R\to R$ heißt ein* **Ring**, *wenn*

1. $(R,+)$ *eine kommutative Gruppe ist – schreibe* $0$ *für das neutrale Element in* $(R,+)$.
2. $\cdot:R\times R\to R$ *assoziativ ist.*
3. *das* **Distributivgesetz** *gilt: Für alle $p,q,r\in R$ folgt:*

$$p\cdot(q+r) = p\cdot q+p\cdot r \tag{3.2.1}$$
$$(p+q)\cdot r = p\cdot r+q\cdot r \tag{3.2.2}$$

*Ein Ring* $(R, +, \cdot)$ *heißt* **kommutativ**, *wenn* $q \cdot r = r \cdot q$ *für alle* $q, r \in R$ *gilt.* $1 \in R$ *heißt* **Einselement**, *wenn* $1 \cdot r = r \cdot 1 = r$ *für alle* $r \in R$ *gilt. Ein kommutativer Ring* $(K, +, \cdot)$ *mit Einselement heißt ein* **Körper** [1], *wenn* $(K \backslash \{0\}, \cdot)$ *eine Gruppe ist.*

**Bemerkung 3.2.2.**

- In einem Ring gilt die **Punkt-vor-Strich-Konvention**, man schreibt also $p \cdot q + p \cdot r$ oder ganz kurz $pq + pr$ statt $(p \cdot q) + (p \cdot r)$.
- In einem Körper $K$ gilt immer $0 \neq 1$, denn 1 ist das neutrale Element von $K \backslash \{0\}$.

**Beispiel 3.2.3.** Auf $\mathbb{Z} = (\mathbb{N} \times \mathbb{N})/_\sim$ – vergleiche Satz 3.1.6 – definiere durch

$$[(k,l)]_\sim \cdot [(m,n)]_\sim := [(k \cdot m + l \cdot n, k \cdot n + l \cdot m)]_\sim$$

eine Multiplikation, dann ist $(\mathbb{Z}, +, \cdot)$ ein kommutativer Ring mit dem Einselement $1 \in \mathbb{N} \subseteq \mathbb{Z}$.

*Beweis.* Ist eine Fleißaufgabe im Überprüfen. □

**Lemma 3.2.4.** *Ein kommutativer Ring* $(R, +, \cdot)$ *mit Einselement ist genau dann ein Körper, wenn jede Gleichung* $aX = b$ *für eine Unbekannte* $X$ *und für alle* $a \in R \backslash \{0\}$ *und* $b \in R$ *genau eine Lösung in* $R$ *besitzt.*

*Beweis.* Nach Definition ist ein kommutativer Ring $R$ mit Einselement $1 \in R$ genau dann ein Körper, wenn $(R \backslash \{0\}, \cdot)$ eine Gruppe ist. Wir erhalten damit:

- Hat jede Gleichung $aX = b$ (mit $a \neq 0$) genau eine Lösung $\lambda(a,b) \in R$, so ist $\lambda(a,1) = a^{-1}$ inverses Element von $a$ in $(R \backslash \{0\}, \cdot)$, also $(R \backslash \{0\}, \cdot)$ eine Gruppe.
- Ist umgekehrt $R$ ein Körper, so ist $a^{-1}b$ offenbar eine Lösung von $aX = b$, die nach den Kürzungsregeln in Lemma 3.1.2 auch eindeutig ist.

□

**Beispiel 3.2.5.** $\mathbb{Z}/2\mathbb{Z} = \{0, 1\}$ wird mit den Verknüpfungen

| $+$ | 0 | 1 |
|---|---|---|
| 0 | 0 | 1 |
| 1 | 1 | 0 |

| $\cdot$ | 0 | 1 |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | 1 |

zu einem Körper. Das Nachprüfen der Rechenregeln geht hier schnell, indem man sich alle möglichen Kombinationen hinschreibt.

**Beispiel* 3.2.6.** Ist $m \in \mathbb{N}$, $m \geq 2$, so läßt sich jede ganze Zahl $z \in \mathbb{Z}$ in eindeutiger Weise in der Form $z = km + r$ mit $k \in \mathbb{Z}$ und $r \in \{0, 1, ..., m-1\}$ schreiben (**Division mit Rest**). Das definiert eine Äquivalenzrelation $\sim$ auf $\mathbb{Z}$ mit den Äquivalenzklassen $\overline{0} := \{km : k \in \mathbb{Z}\}$, $\overline{1} := \{km + 1 : k \in \mathbb{Z}\}$,..., $\overline{m-1} := \{km + (m-1) : k \in \mathbb{Z}\}$. Auf $\mathbb{Z}/m\mathbb{Z} := \mathbb{Z}/_\sim$ wird durch

$$\begin{aligned} \overline{k} + \overline{l} &:= \overline{k+l} \\ \overline{k} \cdot \overline{l} &:= \overline{k \cdot l} \end{aligned}$$

[1] im Englischen: field

eine Addition und Multiplikation definiert, so daß $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$ zu einem kommutativen Ring mit Einselement wird. $\overline{0}$ ist das neutrale Element von $(\mathbb{Z}/m\mathbb{Z}, +)$ und $\overline{1}$ das Einselement. In der Verknüpfungstafel für die Multiplikation in $\mathbb{Z}/4\mathbb{Z}$

| $\cdot$ | $\overline{0}$ | $\overline{1}$ | $\overline{2}$ | $\overline{3}$ |
|---|---|---|---|---|
| $\overline{0}$ | $\overline{0}$ | $\overline{0}$ | $\overline{0}$ | $\overline{0}$ |
| $\overline{1}$ | $\overline{0}$ | $\overline{1}$ | $\overline{2}$ | $\overline{3}$ |
| $\overline{2}$ | $\overline{0}$ | $\overline{2}$ | $\overline{0}$ | $\overline{2}$ |
| $\overline{3}$ | $\overline{0}$ | $\overline{3}$ | $\overline{2}$ | $\overline{1}$ |

sieht man ein weiteres Phänomen, das bei Ringen nicht aber bei Körpern auftreten kann. Wegen $\overline{2} \cdot \overline{2} = \overline{0}$ heißt $\overline{2}$ ein **Nullteiler** in $\mathbb{Z}/4\mathbb{Z}$. Es gilt folgender Satz

**Satz 3.2.7.** *Ist $p \in \mathbb{N}$ eine* **Primzahl** *(dh. es ist $p \geq 2$ und es sind $1$ und $p$ die einzigen natürlichen Zahlen, die $p$ ohne Rest teilen), so ist $\mathbb{Z}/p\mathbb{Z}$ ein Körper.*

**Bemerkung 3.2.8.** In einem Körper $K$ gelten viele der Rechenregeln, über die man nicht weiter nachdenkt, wie etwa: Für alle $a \in K$ gilt:

$$a \cdot 0 = 0 \tag{3.2.3}$$
$$(-1) \cdot a = -a \tag{3.2.4}$$
$$(-1) \cdot (-1) = 1, \tag{3.2.5}$$

Sind $a, b \in K$ mit $ab = 0$, dann folgt $a = 0$ oder $b = 0$

*Beweis.* Zum Beweis muß man sich hier oft nur die genaue Bedeutung der beiden Seiten klar machen: Da $0$ neutrales Element in $(K, +)$ ist, gilt nach dem Distributivgesetz $0 + a \cdot 0 = a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$, woraus $0 = a \cdot 0$ nach der Kürzungsregel in $(K, +)$ folgt. Da $1$ das neutrale Element von $(K \backslash \{0\}, \cdot)$ ist gilt $(-1) \cdot a + a = (-1) \cdot a + 1 \cdot a = (-1+1) \cdot a = 0 \cdot a = 0$, dh. $(-1) \cdot a$ ist inverses Element von $a$ in $(K, +)$, also $(-1) \cdot a = -a$. Wegen $(-1) \cdot (-1) + (-1) = (-1) \cdot (-1) + (-1) \cdot 1 = (-1) \cdot (-1 + 1) = (-1) \cdot 0 = 0$ ist $(-1) \cdot (-1)$ inverses Element von $-1$ in $(K, +)$, also $(-1) \cdot (-1) = 1$. Ist nun $ab = 0$, so dürfen wir oE: $a \neq 0$ annehmen und dann ist $b = 1 \cdot b = a^{-1}(ab) = a^{-1}0 = 0$. $\square$

**Satz 3.2.9.**

1. *Auf $\mathbb{Z} \times (\mathbb{Z} \backslash \{0\})$ definiert*
$$(a, b) \sim (c, d) :\Leftrightarrow ad = bc \tag{3.2.6}$$
*eine Äquivalenzrelation $\sim$. Verwende die vertraute Schreibweise $\frac{a}{b}$ für die Äquivalenzklasse $[(a, b)]_\sim$, dann ist*
$$\mathbb{Q} := (\mathbb{Z} \times (\mathbb{Z} \backslash \{0\}))/_\sim = \left\{ \frac{a}{b} : a \in \mathbb{Z}, b \in \mathbb{Z} \backslash \{0\} \right\}$$

2. *Durch*
$$\frac{a}{b} + \frac{c}{d} := \frac{ad + bc}{bd} \tag{3.2.7}$$
$$\frac{a}{b} \cdot \frac{c}{d} := \frac{ac}{bd} \tag{3.2.8}$$
*werden zwei Verknüpfungen $+ : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ und $\cdot : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ definiert, so daß $(\mathbb{Q}, +, \cdot)$ ein Körper ist.*

3. *Durch die injektive Abbildung* $\begin{array}{rcl} i : \mathbb{Z} & \to & \mathbb{Q} \\ a & \mapsto & \frac{a}{1} \end{array}$ *wird* $\mathbb{Z}$ *identifiziert mit* $i(\mathbb{Z}) \subseteq \mathbb{Q}$ *zu einem (Unter-)Ring in* $\mathbb{Q}$.

*Beweis.* Wegen $ab = ba$ ist $(a,b) \sim (a,b)$, also $\sim$ reflexiv, ferner $(a_1,b_1) \sim (a_2,b_2) \Leftrightarrow a_1b_2 = b_1a_2 \Leftrightarrow (a_2,b_2) \sim (a_1,b_1)$ damit ist $\sim$ symmetrisch. Aus $(a_1,b_1) \sim (a_2,b_2)$ und $(a_2,b_2) \sim (a_3,b_3)$, also $a_1b_2 = b_1a_2$ und $a_2b_3 = b_2a_3$ folgt $a_1b_2a_2b_3 = b_1a_2b_2a_3$ und nach kürzen (im Fall $a_2 \neq 0$) $a_1b_3 = b_1a_3$, dh. $(a_1,b_1) \sim (a_3,b_3)$ und damit ist $\sim$ transitiv und so $\sim$ eine Äquivalenzrelation.

Sind $(a_1,b_1) \sim (a_2,b_2)$ und $(c_1,d_1) \sim (c_2,d_2)$ – also $a_1b_2 = b_1a_2$ und $c_1d_2 = d_1c_2$ – dann gilt

$$\frac{a_1}{b_1} + \frac{c_1}{d_1} = (a_1d_1 + b_1c_1, b_1d_1)_\sim = (a_2d_2 + b_2c_2, b_2d_2)_\sim = \frac{a_2}{b_2} + \frac{c_2}{d_2},$$

denn $(a_1d_1 + b_1c_1)b_2d_2 = (a_1b_2)d_1d_2 + (c_1d_2)b_1b_2 = b_1d_1(a_2d_2 + b_2c_2)$ und

$$\frac{a_1}{b_1} \cdot \frac{c_1}{d_1} = (a_1c_1, b_1d_1)_\sim = (a_2c_2, b_2d_2)_\sim = \frac{a_2}{b_2} \cdot \frac{c_2}{d_2},$$

denn $a_1c_1b_2d_2 = (a_1b_2)(c_1d_2) = b_1d_1a_2c_2$. Damit sind $+$ und $\cdot$ wohldefinierte Abbildungen; da $+$ und $\cdot$ auf $\mathbb{Z}$ kommutativ sind, überträgt sich das auf $\mathbb{Q}$. Für $a,b,c,d,e,f \in \mathbb{Z}$ mit $bdf \neq 0$ ist $\left(\frac{a}{b} + \frac{c}{d}\right) + \frac{e}{f} = \frac{adf + cbf + ebd}{bdf} = \frac{a}{b} + \left(\frac{c}{d} + \frac{e}{f}\right)$ und $\left(\frac{a}{b} \cdot \frac{c}{d}\right) \cdot \frac{e}{f} = \frac{ace}{bdf} = \frac{a}{b} \cdot \left(\frac{c}{d} \cdot \frac{e}{f}\right)$, daher sind $+$ und $\cdot$ assoziativ. Einsetzen in die Definitionen zeigt, daß $0 = \frac{0}{1}$ neutrales Element in $(\mathbb{Q},+)$ und zu $\frac{a}{b} \in \mathbb{Q}$ das inverse Element in $(\mathbb{Q},+)$ durch $\frac{-a}{b}$ gegeben wird. Ebenso folgt durch Einsetzen in die Definition, daß $1 = \frac{1}{1}$ das neutrale Element in $(\mathbb{Q}\backslash\{0\},\cdot)$ ist und $\frac{b}{a}$ das inverse Element zu $\frac{a}{b}$ in $(\mathbb{Q}\backslash\{0\},\cdot)$

Offensichtlich ist $i$ injektiv und $i(1) = 1$, $i(w+z) = i(w)+i(z)$ und $i(w\cdot z) = i(w)\cdot i(z)$ für alle $w,z \in \mathbb{Z}$. $i$ ist daher ein Beispiel für einen **Ringhomomorphismus** und analog zu Lemma 3.1.10 ist $i(\mathbb{Z})$ ein Ring. □

## 3.3 Vektorräume und Untervektorräume

**Definition 3.3.1.** *Es sei* $K$ *ein Körper und für* $\emptyset \neq V$ *gebe es zwei Abbildungen* $+ : V \times V \to V$ *„***Addition***" und* $\cdot : K \times V \to V$ *„***Skalarmultiplikation***".* $V$ *heißt ein* $K$**-Vektorraum***, wenn gilt:*

1. $(V,+)$ *ist eine kommutative Gruppe.*

2. *Für alle* $v,w \in V$ *und* $\lambda,\mu \in K$ *gilt:*

$$(\lambda + \mu) \cdot v = (\lambda \cdot v) + (\mu \cdot v) \tag{3.3.1}$$
$$\lambda \cdot (v + w) = (\lambda \cdot v) + (\lambda \cdot w) \tag{3.3.2}$$
$$(\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v) \tag{3.3.3}$$
$$1 \cdot v = v \tag{3.3.4}$$

*Wir schreiben* $\mathbf{0}$ *für das neutrale Element von* $(V, +)$ *und machen wieder die Punkt-vor-Strich-Konvention.*

**Beispiel 3.3.2.** Ist $K$ ein Körper, $n \in \mathbb{N}$, dann wird durch

$$\begin{aligned} + : K^n \times K^n &\to K^n \\ ((v_1, ..., v_n), (w_1, ..., w_n)) &\mapsto (v_1 + w_1, ..., v_n + w_n) \\ \cdot : K \times K^n &\to K^n \\ (\lambda, (v_1, ..., v_n)) &\mapsto (\lambda v_1, ..., \lambda v_n) \end{aligned}$$

$K^n$ zu einem $K$-Vektorraum. $\mathbf{0} = (0, ..., 0)$ ist das neutrale Element von $(K^n, +)$.

**Lemma 3.3.3.** *Es sei* $V$ *ein* $K$*-Vektorraum, dann gilt für alle* $v \in V$ *und* $\lambda \in K$*:*

$$0 \cdot v = \mathbf{0} \tag{3.3.5}$$

$$\lambda \cdot \mathbf{0} = \mathbf{0} \tag{3.3.6}$$

$$(-1) \cdot v = -v. \tag{3.3.7}$$

*Ist* $\lambda \cdot v = \mathbf{0}$*, so folgt* $\lambda = 0$ *oder* $v = \mathbf{0}$*.*

*Beweis.* $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v$ und nach Subtrahieren von $0 \cdot v$ folgt $\mathbf{0} = (0 \cdot v)$. $\lambda \cdot \mathbf{0} = \lambda \cdot (\mathbf{0} + \mathbf{0}) = \lambda \cdot \mathbf{0} + \lambda \cdot \mathbf{0}$ und nach Subtrahieren von $\lambda \cdot \mathbf{0}$ folgt $\mathbf{0} = \lambda \cdot \mathbf{0}$. $v + (-1) \cdot v = (1 + (-1)) \cdot v = 0 \cdot v = \mathbf{0}$, dh. $(-1) \cdot v$ ist das inverse Element $-v$ zu $v$ in $(V, +)$. Ist $\lambda \cdot v = \mathbf{0}$ und $\lambda = 0$, so gilt die Behauptung, also dürfen wir oE. $\lambda \neq 0$ annehmen. Dann existiert $\lambda^{-1} \in K$, also folgt $v = 1 \cdot v = (\lambda^{-1}\lambda) \cdot v = \lambda^{-1} \cdot (\lambda \cdot v) = \lambda^{-1} \cdot \mathbf{0} = \mathbf{0}$. □

**Definition 3.3.4.** *Ist* $V$ *ein* $K$*-Vektorraum,* $W \subseteq V$ *heißt* $W$ *ein Unterverktorraum, wenn*

- $\mathbf{0} \in W$
- *Für alle* $w_1, w_2 \in W$ *und* $\lambda_1, \lambda_2 \in K$ *gilt:* $\lambda_1 w_1 + \lambda_2 w_2 \in W$*.*

**Bemerkung 3.3.5.** Die Bedingungen in der Definition eines Untervektorraum $W$ eines $K-$Vektorraums $V$ erlauben die Einschränkung der Addition und Skalarmultiplikation $+ : V \times V \to V$ und $\cdot : K \times V \to V$, zu einer Addition $+ : W \times W \to W$ und Skalarmultiplikation $\cdot : K \times W \to W$. Dann bildet $W$ mit dieser von $V$ „geerbten“ Addition und Skalarmultiplikation wieder einen $K-$Vektorraum.

**Lemma 3.3.6.** *Es sei* $V$ *ein* $K$*-Vektorraum und* $W_i, i \in I$ *seien Untervektorräume von* $V$*. Dann ist der Durchschnitt* $\bigcap_{i \in I} W_i$ *ein Untervektorraum von* $V$*.*

*Beweis.* Nach Definition 3.3.4 ist $\mathbf{0} \in W_i$ für alle $i \in I$, also auch $\mathbf{0} \in \bigcap_{i \in I} W_i$. Sind nun $w_1, w_2 \in \bigcap_{i \in I} W_i$ und $\lambda_1, \lambda_2 \in K$, also $w_1, w_2 \in W_i$ für alle $i \in I$, so folgt $\lambda_1 w_1 + \lambda_2 w_2 \in W_i$ für alle $i \in I$, dh. $\lambda_1 w_1 + \lambda_2 w_2 \in \bigcap_{i \in I} W_i$. Nach Definition 3.3.4 ist $\bigcap_{i \in I} W_i$ ein Untervektorraum von $V$. □

**Definition 3.3.7.** *Es sei $V$ ein $K$-Vektorraum und $\emptyset \neq X \subseteq V$, dann heißt*

$$\begin{aligned} \mathrm{lin}(X) \quad := \quad \{v \in V: \ & \textit{Es gibt } n \in \mathbb{N}, x_1, ..., x_n \in X \textit{ und } \lambda_1, ..., \lambda_n \in K \textit{ mit} \\ & v = \lambda_1 x_1 + ... + \lambda_n x_n\} \end{aligned} \tag{3.3.8}$$

*die* **lineare Hülle** *von $X$.*

**Lemma 3.3.8.** *Ist $V$ ein $K$-Vektorraum und $\emptyset \neq X \subseteq V$, so ist $\mathrm{lin}(X)$ ein Untervektorraum von $V$; genauer*

$$\mathrm{lin}(X) = \bigcap_{\substack{W \subseteq V \text{ Untervektorraum} \\ X \subseteq W}} W \tag{3.3.9}$$

*ist der kleinste Untervektorraum von $V$, der $X$ enthält.*

*Beweis.* Für $x \in X$ ist $x = 1 \cdot x \in \mathrm{lin}(X)$, also $X \subseteq \mathrm{lin}(X)$. Nach Lemma 3.3.3 ist $\mathbf{0} = 0 \cdot x \in \mathrm{lin}(X)$ und zu $\lambda, \mu \in K$, $v = \lambda_1 x_1 + ... + \lambda_n x_n \in \mathrm{lin}(X)$ (mit $x_1, ..., x_n \in X$) und $w = \mu_1 y_1 + ... + \mu_m y_m \in \mathrm{lin}(X)$ (mit $y_1, ..., y_m \in X$) ist auch $\lambda v + \mu w = (\lambda\lambda_1)x_1 + ... + (\lambda\lambda_n)x_n + (\mu\mu_1)y_1 + ... + (\mu\mu_m)y_m \in \mathrm{lin}(X)$, also $\mathrm{lin}(X)$ nach Definition 3.3.4 ein Untervektorraum von $V$. Ist $W \subseteq V$ ein Untervektorraum von $V$ mit $X \subseteq W$, so gilt $\lambda_1 x_1 + ... + \lambda_n x_n \in W$ für alle $n \in \mathbb{N}$, $x_1, ..., x_n \in X$ und $\lambda_1, ..., \lambda_n \in K$, also $\mathrm{lin}(X) \subseteq W$. Da $\mathrm{lin}(X)$ ein Untervektorraum ist, der $X$ enthält folgt dann (3.3.9). □

## 3.4 Geordnete Körper

**Definition 3.4.1.** *Es sei $K$ ein Körper und $\leq$ eine Ordnungsrelation auf $K$. $(K, \leq)$ heißt ein* **geordneter Körper**, *wenn gilt:*

- *$\leq$ ist eine totale Ordnung auf $K$.*
- *Sind $p, q, r \in K$ mit $p \leq q$, so folgt $p + r \leq q + r$.*
- *Sind $p, q \in K$ mit $0 \leq p$ und $0 \leq q$, so folgt $0 \leq pq$.*

*Ist $K$ ein geordneter Körper und $a, b \in K$ mit $a < b$, dann heißt*

- *$[a, b] := \{p \in K : a \leq p \leq b\}$ das* **abgeschlossene Intervall** *von $a$ bis $b$.*
- *$]a, b[ := \{p \in K : a < p < b\}$ das* **offene Intervall** *von $a$ bis $b$.*

**Lemma 3.4.2.** *Ist $(K, \leq)$ ein geordneter Körper und $a, b \in K$, dann gilt:*

*a) Aus $a \leq 0$ und $b \geq 0$ folgt $ab \leq 0$.*

*b) Aus $a \leq 0$ und $b \leq 0$ folgt $ab \geq 0$.*

*c) $a^2 \geq 0$.*

*d) Sind $a, b \in K \backslash \{0\}$ dann sind äquivalent: $0 < a \leq b$ und $0 < \dfrac{1}{b} \leq \dfrac{1}{a}$.*

*Beweis.*

a) Zu $a \leq 0$ ist [2] $0 \leq -a$, also nach Definition $(-a) \cdot b = -(a \cdot b) \geq 0$, dh. $ab \leq 0$.

b) Zu $a \leq 0$ und $b \leq 0$ ist $0 \leq -a$ und $0 \leq -b$, also nach Definition $(-a)\cdot(-b) = ab \geq 0$.

c) $a^2 \geq 0$ nach Definition des geordneten Körpers im Fall $a \geq 0$ und $a^2 \geq 0$ als Spezialfall von b) im Fall $a \leq 0$.

d) Ist $a > 0$, so ist auch $\frac{1}{a} > 0$, denn sonst ist $1 = a \cdot \frac{1}{a} \leq 0$ nach a) im Widerspruch zu $1 = 1^2 \geq 0$ und $1 \neq 0$ nach Bemerkung 3.2.2.
„$\Rightarrow$“: Es sei $0 < a \leq b$, dann ist $ab > 0$ und $b - a \geq 0$ also $\frac{1}{a} - \frac{1}{b} = \frac{1}{ab}(b-a) \geq 0$, also $\frac{1}{b} \leq \frac{1}{a}$ und $0 < \frac{1}{b}$ und $0 < \frac{1}{a}$ nach der Vorbemerkung, dh. $0 < \frac{1}{b} \leq \frac{1}{a}$.
„$\Leftarrow$“: Es sei $0 < \frac{1}{b} \leq \frac{1}{a}$ dann ist nach der Vorbemerkung $b > 0$ und $a > 0$ und $\frac{1}{a} - \frac{1}{b} = \frac{1}{ab}(b-a) \geq 0$. Wegen $\frac{1}{ab} = \frac{1}{a} \cdot \frac{1}{b} > 0$ ist also $b - a \geq 0$. □

**Lemma 3.4.3.** *Es sei $(K, \leq)$ ein geordneter Körper und $n \in \mathbb{N}$, $K_+ := \{a \in K : a > 0\}$, dann ist* $\begin{array}{rcl} x^n : K_+ \cup \{0\} & \to & K_+ \cup \{0\} \\ x & \mapsto & x^n \end{array}$ *streng monoton steigend.*

*Beweis.* Zu $a, b \in K$ ist

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + ... + a^2b^{n-3} + ab^{n-2} + b^{n-1}),$$

wie man durch ausmultiplizieren sieht. Ist nun $0 \leq a < b$, dann gilt $a - b < 0$ und $a^{n-1}, a^{n-2}b, ..., ab^{n-2}, b^{n-1} \geq 0$, also $a^{n-1} + a^{n-2}b + ... + ab^{n-2} + b^{n-1} > 0$ und damit

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + ... + a^2b^{n-3} + ab^{n-2} + b^{n-1}) < 0,$$

also $x^n$ streng monoton steigend. □

**Satz 3.4.4.** $\mathbb{Q}_+ := \{\frac{a}{b} \in \mathbb{Q} : a, b \in \mathbb{N}\}$ *heißt die Menge der positiven rationalen Zahlen,* $\mathbb{Q}_- := \{-\frac{a}{b} \in \mathbb{Q} : a, b \in \mathbb{N}\}$ *heißt die Menge der negativen rationalen Zahlen.* $\mathbb{Q}_+$, $\{0\}$ *und* $\mathbb{Q}_-$ *bilden eine Zerlegung von* $\mathbb{Q}$ *und durch*

$$r \leq s :\Leftrightarrow s - r \in \mathbb{Q}_+ \cup \{0\} \tag{3.4.10}$$

*wird eine totale Ordnung $\leq$ auf $\mathbb{Q}$ definiert, so daß $(\mathbb{Q}, +, \cdot, \leq)$ ein geordneter Körper ist.*

*Beweis.* Sind $a, b \in \mathbb{N}$, dann gilt nach der Definition der Brüche $\frac{a}{-b} = \frac{-a}{b} = -\frac{a}{b} \neq 0$ und $\frac{a}{b} = \frac{-a}{-b} \neq 0$ und damit bekommt man die Zerlegung von $\mathbb{Q}$ in paarweise disjunkte Mengen $\mathbb{Q} = \mathbb{Q}_+ \cup \{0\} \cup \mathbb{Q}_-$. Sind $r, s \in \mathbb{Q}$, dann ist auch $s - r \in \mathbb{Q}$ und es gibt daher nur die beiden Möglichkeiten $s - r \in \mathbb{Q}_+ \cup \{0\}$ oder $s - r \in \mathbb{Q}_-$, also $r - s \in \mathbb{Q}_+$ und damit $r \leq s$ oder $s \leq r$ und daher ist $\mathbb{Q}$ total geordnet. Sind $p, q, r \in \mathbb{Q}$ mit $p \leq q$ – also $q - p \in \mathbb{Q}_+ \cup \{0\}$ – so ist auch $p + r \leq q + r$ –dh. $(q + r) - (p + r) = q - p \in \mathbb{Q}_+ \cup \{0\}$. Sind $p = \frac{a}{b}, q = \frac{c}{d} \in \mathbb{Q}_+$ mit $a, b, c, d \in \mathbb{N}$, dann ist $pq = \frac{ac}{bd} \in \mathbb{Q}_+$ und damit ist $(\mathbb{Q}, \leq)$ ein geordneter Körper. □

[2] für $a = 0$ gilt das und für $a < 0$ führt $-a < 0$ zum Widerspruch $0 = a + (-a) < 0 + (-a) = -a \leq 0$

**Lemma 3.4.5.** *Ist $(K, \leq)$ ein geordneter Körper, dann gibt es einen monoton steigenden injektiven Ringhomomorphismus $i : \mathbb{Q} \to K$.*

**Bemerkung 3.4.6.**

- Lemma 3.4.5 zeigt uns, daß die Körper $\mathbb{Z}/p\mathbb{Z}$ nicht zu geordneten Körpern gemacht werden können.
- In der Situation von Lemma 3.4.5 ist $i(\mathbb{Q}) \subseteq K$ ein Körper – den man dann üblicherweise wieder mit $\mathbb{Q}$ identifiziert. Insbesondere hat man in dieser Situation die natürlichen Zahlen mit einer Teilmenge des geordneten Körpers $(K, \leq)$ identifiziert.

**Definition 3.4.7.** *Ein geordneter Körper $(K, \leq)$ heißt* **archimedisch** *geordnet, wenn es für alle $p, q \in K$ mit $p > 0$ und $q > 0$ ein $n \in \mathbb{N}$ gibt mit $p < nq$.*

**Lemma 3.4.8.** $(\mathbb{Q}, \leq)$ *ist ein archimedisch geordneter Körper.*

*Beweis.* Sind $r, s \in \mathbb{Q}_+$, dann gibt es $a, b, c \in \mathbb{N}$ mit $s = \frac{a}{b}$ und $r = \frac{c}{b}$. Für den Nachweis von $s < nr$ mit $n \in \mathbb{N}$ genügt es also ein $n \in \mathbb{N}$ mit $a < nc$ zu finden. $n = a + 1$ erfüllt diese Bedingung. $\square$

# Kapitel 4

# Reelle und komplexe Zahlen

## 4.1 Reelle Zahlen $\mathbb{R}$

**Definition 4.1.1.** *Der Körper $\mathbb{R}$ der reellen Zahlen ist ein archimedisch geordneter Körper $(\mathbb{R}, +, \cdot, \leq)$ für den das* **Intervallschachtelungsaxiom** *gilt: Ist $([a_n, b_n])_{n\in\mathbb{N}}$ eine Folge von abgeschlossenen Intervallen, die für alle $n \in \mathbb{N}$ die Bedingungen*

$$a_n \leq b_n,\ a_n \leq a_{n+1} \quad \textit{und} \quad b_{n+1} \leq b_n$$

*erfüllen, dann gilt*

$$\bigcap_{n\in\mathbb{N}} [a_n, b_n] \neq \emptyset. \tag{4.1.1}$$

**Lemma 4.1.2.** *Ist $x \in \mathbb{R}$ mit $x > 0$, so gibt es $n \in \mathbb{N}$ mit $\frac{1}{n} < x$. Insbesondere gibt es $k \in \mathbb{N}$ mit $\frac{1}{2^k} < x$.*

*Beweis.* Nach dem archimedischen Axiom (angewandt auf 1 und $x$) gibt es ein $n \in \mathbb{N}$ mit $1 < nx$, also $\frac{1}{n} < x$. Für alle $k \in \mathbb{N}$ ist $2^k > k$, wie wir durch Induktion zeigen:

- Induktionsanfang: $k = 1$: $2^1 = 2 > 1$.
- Induktionsschritt $k \to k+1$: $2^{k+1} = 2 \cdot 2^k > 2 \cdot k \geq k+1$.

Damit ist $\frac{1}{2^n} < \frac{1}{n} < x$. □

**Lemma 4.1.3.** *Sind $a, b \in \mathbb{R}$ mit $a < b$, dann gibt es $q \in \mathbb{Q} \cap ]a, b[$.*

*Beweis.* Wir betrachten oE. den Fall $b > 0$ (bei $b \leq 0$ betrachte $]-b, -a[$); es ist $b - a > 0$ und nach Lemma 4.1.2 gibt es $n \in \mathbb{N}$ mit $\frac{1}{n} < b - a$. Nach dem archimedischen Axiom gibt es ferner ein $k \in \mathbb{N}$ mit $b < k \cdot \frac{1}{n}$. Nach Satz 2.1.12 existiert $h := \min\{k \in \mathbb{N} : b \leq k\frac{1}{n}\}$, da $\{k \in \mathbb{N} : b \leq k\frac{1}{n}\} \neq \emptyset$ und dann ist $\frac{h-1}{n} < b$. Ferner ist $a < \frac{h-1}{n}$, denn sonst bleibt nur $a \geq \dfrac{h-1}{n}$ und damit $\dfrac{h-1}{n} \leq a < b \leq \dfrac{h}{n}$, also $b - a \leq \frac{1}{n}$ im Widerspruch zur Wahl von $n$ mit $\dfrac{1}{n} < b - a$. Wir haben damit $\dfrac{h-1}{n} \in \mathbb{Q} \cap ]a, b[$ gezeigt. □

**Satz 4.1.4.** *Ist* $\emptyset \neq X \subseteq \mathbb{R}$ *nach oben beschränkt, so existiert* $\sup X$ *in* $\mathbb{R}$.

*Beweis.* Es sei $a \in X$ und $b \in \mathbb{R}$ eine obere Schranke von $X$; oE. ist $a < b$, denn für $b \in X$ gilt $b = \max(X) = \sup(X)$. Dann gibt es nach dem archimedischen Axiom zu jedem $n \in \mathbb{N}_0$ ein $m \in \mathbb{N}$ mit $b - a \leq m \cdot 2^{-n}$. Daher ist für alle $n \in \mathbb{N}$

$$Y_n := \{m \in \mathbb{N} : a + m \cdot 2^{-n} \text{ obere Schranke von } X\} \neq \emptyset$$

und damit gibt es nach Satz 2.1.12 ein kleinstes Element $p_n = \min(Y_n) \in Y_n$. Definiere

$$I_n := [a + (p_n - 1) \cdot 2^{-n}, a + p_n \cdot 2^{-n}].$$

Da $a + (p_n - 1) \cdot 2^{-n}$ keine obere Schranke von $X$ ist gilt $I_n \cap X \neq \emptyset$ und außerdem ist $p_{n+1} = 2p_n$ oder $p_{n+1} = 2p_n - 1$. [1] Daher ist $I_{n+1} \subseteq I_n$. Nach dem Intervallschachtelungsaxiom gilt $\bigcap_{n\in\mathbb{N}} I_n \neq \emptyset$. Wir zeigen nun sukzessive:

- $\bigcap_{n\in\mathbb{N}} I_n = \{\gamma\}$,
  denn wenn es $\alpha, \beta \in \bigcap_{n\in\mathbb{N}} I_n$ mit $\alpha < \beta$ gibt, so gilt $[\alpha, \beta] \subseteq I_n$, dh. $\beta - \alpha \leq 2^{-n}$ für alle $n \in \mathbb{N}$. Damit folgt $n(\beta - \alpha) < 2^n(\beta - \alpha) \leq 1$ für alle $n \in \mathbb{N}$, was dem archimedischen Axiom widerspricht – wonach es $k \in \mathbb{N}$ mit $1 < k(\beta - \alpha)$ gibt.

- $\gamma$ ist obere Schranke von $X$,
  denn sonst gibt es ein $x \in X$ mit $\gamma < x$, daher gibt es nach Lemma 4.1.2 ein $n \in \mathbb{N}$ mit $2^{-n} < x - \gamma$ und wegen $\gamma \in I_n$ ist

  $$a + p_n \cdot 2^{-n} = 2^{-n} + (a + (p_n - 1) \cdot 2^{-n}) \leq 2^{-n} + \gamma < x,$$

  im Widerspruch zur Wahl von $p_n$, wonach $a + p_n 2^{-n}$ obere Schranke von $X$ ist.

- $\gamma = \sup X$,
  denn ist $y$ obere Schranke von $X$ mit $\gamma - y > 0$, so gibt es nach Lemma 4.1.2 ein $n \in \mathbb{N}$ mit $2^{-n} < \gamma - y$; aus $\gamma \in I_n$ folgt:

  $$y < \gamma - 2^{-n} \leq a + p_n \cdot 2^{-n} - 2^{-n} = a + (p_n - 1) \cdot 2^{-n}$$

  im Widerspruch zu $I_n \cap X \neq \emptyset$. □

**Lemma 4.1.5.** *Ist* $\emptyset \neq X \subseteq \mathbb{R}$ *nach oben beschränkt, dann ist* $\gamma := \sup X$ *die eindeutig bestimmte reelle Zahl mit den Eigenschaften:*

---

[1] Da $a + (p_n - 1)2^{-n}$ keine obere Schranke von $X$ ist, gibt es ein $x \in X$ mit $x > a + (p_n - 1)2^{-n}$. Gibt es ein $x \in X$ mit $x > a + (p_n - 1)2^{-n} + 2^{-n-1}$, so ist $p_{n+1} = 2p_n$. Wenn es kein $x \in X$ mit $x > a + (p_n - 1)2^{-n} + 2^{-n-1}$ gibt, dann ist $p_{n+1} = 2p_n - 1$.

*a)* $\gamma$ *ist obere Schranke von* $X$.

*b) Für jedes* $n \in \mathbb{N}$ *existiert ein* $x_n \in X$ *mit*

$$\gamma - \frac{1}{n} < x_n \leq \gamma. \tag{4.1.2}$$

*Beweis.* Ist $\emptyset \neq X \subseteq \mathbb{R}$ nach oben beschränkt, dann existiert $\gamma := \sup(X)$ nach Satz 4.1.4. Da $\gamma$ die kleinste obere Schranke von $X$ ist, gilt die Bedingung a) und für jedes $\varepsilon > 0$ ist $\gamma - \varepsilon$ keine obere Schranke von $X$, dh. es gibt ein $x(\varepsilon) \in X$ mit $\gamma - \varepsilon < x(\varepsilon)$. Für die Wahl von $\varepsilon = \frac{1}{n}$ finden wir also $x_n \in X$ mit $\gamma - \frac{1}{n} < x_n \leq \gamma$, also ist auch b) erfüllt. Sei nun $\gamma \in \mathbb{R}$ eine Zahl, die a) und b) erfüllt und $y$ eine obere Schranke von $X$, dann ist $y \geq x_n$ für jedes $n \in \mathbb{N}$, also $\gamma - \frac{1}{n} \leq y$ für alle $n \in \mathbb{N}$, oder $y \geq \gamma$ und daher ist $\gamma = \sup(X)$ die kleinste obere Schranke von $X$. □

**Lemma 4.1.6.** *Ist* $\emptyset \neq X \subseteq \mathbb{R}$ *nach oben beschränkt, dann ist* $-X := \{-x : x \in X\}$ *nach unten beschränkt und es gilt:* $\inf(-X) = -\sup(X)$

*Beweis.* Es sei $\beta \in \mathbb{R}$. Nach Definition ist $-\beta$ genau dann obere Schranke von $X$, wenn $\beta$ untere Schranke von $-X$ ist, also ist insbesondere $-X$ nach unten beschränkt. Nach Satz 4.1.4 existiert $\gamma := \sup(X)$ und es gilt $\gamma \leq (-\beta)$ für jede obere Schranke $-\beta$ von $X$, also $-\gamma \geq \beta$ für jede untere Schranke $\beta$ von $-X$. Daher ist $\inf(-X) = -\sup(X)$. □

**Korollar 4.1.7.** *Ist* $\emptyset \neq X \subseteq \mathbb{R}$ *nach unten beschränkt, so existiert* $\inf X$ *in* $\mathbb{R}$ *und* $\gamma := \inf X$ *ist die reelle Zahl mit den Eigenschaften:*

*a)* $\gamma$ *ist untere Schranke von* $X$.

*b) Für jedes* $n \in \mathbb{N}$ *existiert ein* $x_n \in X$ *mit*

$$\gamma \leq x_n < \gamma + \frac{1}{n}. \tag{4.1.3}$$

## 4.2 Komplexe Zahlen $\mathbb{C}$

**Definition 4.2.1.** *Wir schreiben* $1$ *für* $(1,0) \in \mathbb{R}^2$ *und* $i$ *für* $(0,1) \in \mathbb{R}^2$ *und* $x + iy$ *für* $(x,y) \in \mathbb{R}^2$ *und* $\mathbb{C}$ *statt* $\mathbb{R}^2$, *wenn wir auf* $\mathbb{C}$ *folgendermaßen Addition und Multiplikation definieren:*

$$\begin{aligned} + : \mathbb{C} \times \mathbb{C} &\to \mathbb{C} \\ (x_1 + iy_1, x_2 + iy_2) &\mapsto x_1 + x_2 + i(y_1 + y_2) \end{aligned} \tag{4.2.1}$$

$$\begin{aligned} \cdot : \mathbb{C} \times \mathbb{C} &\to \mathbb{C} \\ (x_1 + iy_1, x_2 + iy_2) &\mapsto x_1x_2 - y_1y_2 + i(x_1y_2 + x_2y_1) \end{aligned} \tag{4.2.2}$$

**Lemma 4.2.2.** $(\mathbb{C}, +, \cdot)$ *ist ein Körper,* $0 = 0 + i0$ *und* $1 = 1 + i0$ *sind die neutralen Elemente von* $(\mathbb{C}, +)$ *bzw.* $(\mathbb{C}\backslash\{0\}, \cdot)$. *Zu jedem* $x + iy \in \mathbb{C}\backslash\{0\}$ *ist* $\dfrac{x - iy}{x^2 + y^2}$ *das inverse Element in* $(\mathbb{C}\backslash\{0\}, \cdot)$. $\{x + i0 : x \in \mathbb{R}\} \subseteq \mathbb{C}$ *ist ein zu* $\mathbb{R}$ *isomorpher Unterkörper von* $\mathbb{C}$.

*Beweis.* Aufgabe □

**Lemma 4.2.3.** *$i^2 = -1$ und $\mathbb{C}$ kann nicht zu einem geordneten Körper gemacht werden.*

*Beweis.* Nach Definition der Multiplikation in $\mathbb{C}$ gilt: $i^2 = (0+i)^2 = -1$. In einem geordneten Körper $K$ gilt $a^2 \geq 0$ für alle $a \in K$ nach Lemma 3.4.2, daher kann $\mathbb{C}$ kein geordneter Körper sein. □

**Definition 4.2.4.** *Zu einer komplexen Zahl $z = x + iy \in \mathbb{C}$ heißt*

- *$\bar{z} := x - iy$ die zu $z$* **komplex konjugierte Zahl**
- *$x$ der* **Realteil** *von $z = x + iy$; Schreibweise $x = \mathrm{Re}\,(x+iy)$*
- *$y$ der* **Imaginärteil** *von $z = x + iy$; Schreibweise $y = \mathrm{Im}\,(x+iy)$*

**Lemma 4.2.5.** *Für alle $z \in \mathbb{C}$ gilt:*

$$\mathrm{Re}\, z = \frac{1}{2}(z + \bar{z}) \qquad \mathrm{Im}\, z = \frac{1}{2i}(z - \bar{z}) \qquad z\bar{z} \geq 0 \qquad \bar{\bar{z}} = z \tag{4.2.3}$$

*Beweis.* Ist $z = x + iy = \mathrm{Re}\, z + i\mathrm{Im}\, z$, so ist

- $\bar{z} = x - iy$, also $z + \bar{z} = x + iy + (x - iy) = 2x$, dh. $x = \mathrm{Re}\, z = \dfrac{1}{2}(z + \bar{z})$.
- $z - \bar{z} = x + iy - (x - iy) = 2iy$, dh. $y = \mathrm{Im}\, z = \dfrac{1}{2i}(z - \bar{z})$.
- $z\bar{z} = (x+iy)(x-iy) = x^2 + y^2 + i(xy - xy) = x^2 + y^2 \geq 0$.
- $\bar{\bar{z}} = \overline{x - iy} = x + iy = z$. □

# Kapitel 5

# Konvergenz von Folgen

## 5.1 Konvergenz in metrischen Räumen

**Definition 5.1.1.** *Sei $X$ eine Menge, $d : X \times X \to [0,\infty[$. $d$ heißt* **Metrik** *auf $X$ und $(X,d)$ heißt* **metrischer Raum***, wenn für alle $x, y, z \in X$ gilt:*

$$d(x,y) = 0 \text{ genau dann wenn } x = y \tag{5.1.1}$$
$$d(x,y) = d(y,x) \tag{5.1.2}$$
$$d(x,y) \leq d(x,z) + d(z,y) \qquad \text{(Dreiecksungleichung)} \tag{5.1.3}$$

**Beispiel 5.1.2.**

a) Mit $\begin{array}{rcl} d : \mathbb{R} \times \mathbb{R} & \to & [0,\infty[ \\ (x,y) & \mapsto & |x-y| \end{array}$ wird $(\mathbb{R}, d)$ zu einem metrischen Raum. [1]

b) Auch jede Teilmenge von $\mathbb{R}$ wird durch Einschränkung dieser Metrik zu einem metrischen Raum, zB. $\begin{array}{rcl} d : \mathbb{Q} \times \mathbb{Q} & \to & \mathbb{Q} \\ (x,y) & \mapsto & |x-y| \end{array}$.

c) Für $\emptyset \neq X$ ist $\begin{array}{rcl} d : X \times X & \to & [0,\infty[ \\ (x,y) & \mapsto & \begin{cases} 0 & \text{für } x = y \\ 1 & \text{für } x \neq y \end{cases} \end{array}$ eine Metrik auf $X$.

**Definition 5.1.3.**

- *Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum $(X,d)$ heißt* **Cauchyfolge***, wenn es zu jedem $\varepsilon > 0$ ein $k = k(\varepsilon) \in \mathbb{N}$ gibt mit*

$$d(x_m, x_n) < \varepsilon \quad \text{für alle } m, n \geq k. \tag{5.1.4}$$

[1] Die Dreiecksungleichung erhält man durch Fallunterscheidung:
Für $x = y$ ist $d(x,y) = 0 \leq d(x,z) + d(y,z)$. Es sei also oE. $x < y$, dann ist

$$d(x,z) + d(z,y) = \left\{ \begin{array}{ll} (x-z) + (y-z) \geq 0 + (y-x) & \text{für } z \leq x \\ (z-x) + (y-z) = y - x & \text{für } z \in ]x,y[ \\ (z-x) + (z-y) \geq (y-x) + (z-y) \geq (y-x) + 0 & \text{für } z \geq y \end{array} \right\}$$

- *Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum $(X,d)$ heißt* **konvergent**, *wenn es ein $x \in X$ gibt, so daß es für alle $\varepsilon > 0$ ein $k = k(\varepsilon) \in \mathbb{N}$ gibt mit*

$$d(x, x_n) < \varepsilon \quad \text{für alle } n \geq k. \tag{5.1.5}$$

*In diesem Fall schreibt man $x = \lim\limits_{n\to\infty} x_n$ und nennt $x$ den* **Grenzwert** *oder* **Limes** *von $(x_n)_{n\in\mathbb{N}}$.*

- *Ist* $\begin{array}{rcl}\varphi : \mathbb{N} & \to & \mathbb{N} \\ n & \mapsto & \varphi(n)\end{array}$ *streng monoton steigend, so heißt $(x_{\varphi(n)})_{n\in\mathbb{N}}$ eine* **Teilfolge** *von $(x_n)_{n\in\mathbb{N}}$.*

**Beispiel 5.1.4.**

- Jede Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum $(X,d)$, die ab einem Index $N \in \mathbb{N}$ konstant ist – die also $x_n = x$ für alle $n \in \mathbb{N}, n \geq N$ erfüllt – ist konvergent mit $x = \lim\limits_{n\to\infty} x_n$.

- Ist $(x_n)_{n\in\mathbb{N}}$ eine Folge, so ist $(x_{2n})_{n\in\mathbb{N}}$ eine Teilfolge davon – die zugehörige streng monoton steigende Funktion ist $\begin{array}{rcl}\varphi : \mathbb{N} & \to & \mathbb{N} \\ n & \mapsto & 2n\end{array}$.

- Nicht jede konvergente Folge $(x_n)_{n\in\mathbb{N}}$ muß ab einem Index $N$ konstante Folgenglieder haben, wie das Beispiel der Folge $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ in $\mathbb{R}$ zeigt. Nach Lemma 4.1.2 gibt es für jedes $\varepsilon > 0$ ein $k = k(\varepsilon) \in \mathbb{N}$ mit $\frac{1}{k} < \varepsilon$, daher gilt für alle $n \geq k$:

$$\left|0 - \frac{1}{n}\right| = \frac{1}{n} \leq \frac{1}{k} < \varepsilon$$

und somit ist $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ konvergent mit $0 = \lim\limits_{n\to\infty} \frac{1}{n}$.

**Lemma 5.1.5.** *Ist $(x_n)_{n\in\mathbb{N}}$ eine konvergente Folge in einem metrischen Raum $(X,d)$, so ist der Grenzwert $x = \lim\limits_{n\to\infty} x_n$ eindeutig.*

*Beweis.* Es sei $(x_n)_{n\in\mathbb{N}}$ eine konvergente Folge und $x, y \in X$ seien Grenzwerte. Für jedes $k \in \mathbb{N}$ gibt es dann $M(k), N(k) \in \mathbb{N}$ mit $d(x, x_n) \leq \frac{1}{k}$ für $n \geq M(k)$ und $d(y, x_n) \leq \frac{1}{k}$ für $n \geq N(k)$. Ist also $n \geq \max\{M(k), N(k)\}$, so gilt: $d(x,y) \leq d(x,x_n) + d(x_n,y) \leq \frac{2}{k}$. Da wir diese Abschätzung für alle $k \in \mathbb{N}$ bekommen gilt $d(x,y) \in \bigcap\limits_{k\in\mathbb{N}} [0, \frac{2}{k}] = \{0\}$, also $d(x,y) = 0$ und daher $x = y$. □

**Lemma 5.1.6.** *Jede konvergente Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum $(X,d)$ ist eine Cauchyfolge.*

*Beweis.* Ist $\lim_{n\to\infty} x_n = x$ und $\varepsilon > 0$, dann gibt es $k = k(\varepsilon) \in \mathbb{N}$ mit $d(x, x_n) < \frac{\varepsilon}{2}$ für alle $n \in \mathbb{N}$ mit $n \geq k$. Nach der Dreiecksungleichung gilt $d(x_m, x_n) \leq d(x_m, x) + d(x, x_n) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ für alle $m, n \geq k$ und daher ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge. □

**Lemma 5.1.7.** *Ist $(x_n)_{n\in\mathbb{N}}$ eine konvergente Folge in einem metrischen Raum $(X, d)$ mit $x = \lim_{n\to\infty} x_n$ und $(x_{\varphi(n)})_{n\in\mathbb{N}}$ eine Teilfolge, dann ist auch $(x_{\varphi(n)})_{n\in\mathbb{N}}$ konvergent mit $x = \lim_{n\to\infty} x_{\varphi(n)}$.*

*Beweis.* Zu $\varepsilon > 0$ sei $N(\varepsilon) \in \mathbb{N}$ so gewählt, daß $d(x, x_n) < \varepsilon$ für alle $n \geq N(\varepsilon)$ gilt. Da $\varphi : \mathbb{N} \to \mathbb{N}$ streng monoton steigend ist, gibt es dann $J(\varepsilon) \in \mathbb{N}$, so daß $\varphi(n) \geq N(\varepsilon)$ für alle $n \geq J(\varepsilon)$ gilt. Damit ist aber $d(x, x_{\varphi(n)}) < \varepsilon$ für alle $n \geq J(\varepsilon)$, dh. $x = \lim_{n\to\infty} x_{\varphi(n)}$. □

**Lemma 5.1.8.** *Ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge, so ist $(x_n)_{n\in\mathbb{N}}$ beschränkt, dh. es gibt ein $x \in X$ und $C \in \mathbb{R}_+$ mit $d(x, x_n) \leq C$ für alle $n \in \mathbb{N}$.*

*Beweis.* Ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge und $\varepsilon > 0$, so gibt es ein $k \in \mathbb{N}$ mit $d(x_m, x_n) < \varepsilon$ für alle $m, n \geq k$, dann ist für $x \in X$ wegen der Dreiecksungleichung $d(x, x_n) \leq d(x, x_k) + d(x_k, x_n)$ und damit

$$\begin{aligned} d(x, x_j) &\leq \max\{d(x, x_1), ..., d(x, x_k), d(x, x_k) + d(x_k, x_{k+1}), d(x, x_k) + d(x_k, x_{k+2}), ...\} \\ &\leq \max\{d(x, x_1), ..., d(x, x_k), d(x, x_k) + \varepsilon\} =: C < \infty \end{aligned}$$

für alle $j \in \mathbb{N}$. □

**Definition 5.1.9.** *Ein metrischer Raum $(X, d)$ heißt* **vollständig***, wenn jede Cauchyfolge in $X$ konvergent ist.*

**Satz 5.1.10.** *Jede Cauchyfolge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ ist eine konvergente Folge in $\mathbb{R}$.*

*Beweis.* Es sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $\mathbb{R}$. Definiere eine Folge $(n_k)_{k\in\mathbb{N}}$ von natürlichen Zahlen $n_k$ in folgender Weise: $n_1 := 1$ und $n_{k+1} := \min M_{k+1}$, wobei

$$M_{k+1} := \{m \in \mathbb{N} : m > n_k;\ \text{für alle } p \geq m \text{ und } q \geq m \text{ gilt: } |x_p - x_q| < 2^{-(k+2)}\}.$$

Diese Definition ist möglich, denn weil $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist, gibt es für die positive Zahl $2^{-(k+2)} > 0$ ein $N(k) \in \mathbb{N}$ mit $|x_p - x_q| < 2^{-(k+2)}$ für alle $p, q \geq N(k)$. Damit sind alle $m \in \mathbb{N}$ mit $m > n_k$ und $m \geq N(k)$ in $M_{k+1}$ enthalten. Damit ist $M_{k+1} \neq \emptyset$, also existiert $n_{k+1}$ nach Satz 2.1.12. und $M_{k+1} = \{m \in \mathbb{N} : m \geq n_{k+1}\}$ und die Folge $(n_k)_{k\in\mathbb{N}}$ ist streng monoton steigend. Zu $k \in \mathbb{N}$ sei $I_k$ das abgeschlossene Intervall

$$I_k := [x_{n_k} - 2^{-k}, x_{n_k} + 2^{-k}].$$

Nach obiger Konstruktion ist $n_{k+1} \in M_{k+1} \subseteq M_k$, $k \geq 2$ also $|x_{n_k} - x_{n_{k+1}}| < 2^{-k-1}$ nach Definition von $M_k$ und damit gilt

$$x_{n_{k+1}} - 2^{-k-1} \geq x_{n_k} - |x_{n_k} - x_{n_{k+1}}| - 2^{-k-1} > x_{n_k} - 2^{-k-1} - 2^{-k-1} = x_{n_k} - 2^{-k}$$

für die linke Intervallgrenze $x_{n_{k+1}} - 2^{-k-1}$ von $I_{k+1}$ und die linke Intervallgrenze $x_{n_k} - 2^{-k}$ von $I_k$. Ebenso folgt

$$x_{n_{k+1}} + 2^{-k-1} \leq x_{n_k} + |x_{n_k} - x_{n_{k+1}}| + 2^{-k-1} \leq x_{n_k} + 2^{-k-1} + 2^{-k-1} = x_{n_k} + 2^{-k}$$

für die rechten Intervallgrenzen von $I_{k+1}$ und $I_k$, was $I_{k+1} \subseteq I_k$ zeigt. Nach dem Intervallschachtelungsaxiom ist $\bigcap_{n\in\mathbb{N}} I_n \neq \emptyset$. Die Intervallänge von $I_k$ ist $2^{-k+1}$ und wegen $0 = \lim_{k\to\infty} 2^{-k+1}$ erhalten wir wie beim Beweis von Satz 4.1.4 ein $x \in \mathbb{R}$ mit $\bigcap_{n\in\mathbb{N}} I_n = \{x\}$. Für jedes $m \in \mathbb{N}$ mit $m \geq n_k$ gilt $x_m \in I_k$, also $|x - x_m| \leq 2^{-k+1}$, was $x = \lim_{n\to\infty} x_n$ nach Lemma 4.1.2 zeigt. □

**Bemerkung 5.1.11.** Wir werden im nächsten Abschnitt sehen, daß $\mathbb{Q}$ nicht vollständig ist. Wir konstruieren dort eine Cauchyfolge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in \mathbb{Q}$, so daß der Limes - der nach Satz 5.1.10 ja existiert, wenn wir $(x_n)_{n\in\mathbb{N}}$ als reelle Folge betrachten – die Bedingung $x = \lim_{n\to\infty} x_n \in \mathbb{R}\backslash\mathbb{Q}$ erfüllt. Eine solche Cauchyfolge konvergiert wegen der Eindeutigkeit des Grenzwerts nicht in $\mathbb{Q}$, da der Grenzwert $x \notin \mathbb{Q}$.
Die Vollständigkeit ist eine erstrebenswerte Eigenschaft; wenn man etwa Näherungslösungen einer gegebenen Gleichung mit einer geeigneten Fehlerabschätzung hat, bilden diese eine Cauchyfolge und dann wäre es ja schön, wenn diese Näherungslösungen gegen einen Grenzwert konvergieren, der die gegebene Gleichung löst.

**Definition 5.1.12.** *Es sei $(X, d)$ ein metrischer Raum und $(x_n)_{n\in\mathbb{N}}$ eine Folge in $X$, dann heißt $x \in X$ ein* **Häufungswert** *(oder Häufungspunkt) von $(x_n)_{n\in\mathbb{N}}$, wenn es eine Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ gibt mit $x = \lim_{n\to\infty} x_{\varphi(n)}$.*

**Beispiel 5.1.13.**

a) Ist $(x_n)_{n\in\mathbb{N}}$ eine konvergente Folge mit $x = \lim_{n\to\infty} x_n$, dann ist $x$ der einzige Häufungswert von $(x_n)_{n\in\mathbb{N}}$, denn nach Lemma 5.1.7 konvergiert jede Teilfolge $(x_{\varphi(n)})_{n\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ gegen $x$.

b) Die Folge $((-1)^n)_{n\in\mathbb{N}}$ ist nicht konvergent, aber 1 und $-1$ sind alle Häufungswerte von $((-1)^n)_{n\in\mathbb{N}}$. Denn offenbar ist $((-1)^{2n})_{n\in\mathbb{N}}$ eine Teilfolge von $((-1)^n)_{n\in\mathbb{N}}$, die konstant nur den Wert 1 annimmt und ebenso ist $((-1)^{2n+1})_{n\in\mathbb{N}}$ eine Teilfolge, die konstant ist – und zwar gleich $-1$. Daher sind $-1$ und 1 Häufungswerte. Ist $x \in \mathbb{R}\backslash\{-1, 1\}$, dann ist $\delta := \frac{1}{2}\min\{|x - 1|, |x + 1|\} > 0$ und so gibt es kein Folgenglied $(-1)^n$ mit $|x-(-1)^n| < \delta$, also ist $x$ kein Häufungswert von $((-1)^n)_{n\in\mathbb{N}}$.

c) Eine Folge $(x_n)_{n\in\mathbb{N}}$, die gerade die Häufungswerte $0, 1$ und 2 besitzt kann man etwa durch

$$x_n = \begin{cases} 0 & \text{für } n = 3k \\ 1 & \text{für } n = 3k+1 \\ 2 & \text{für } n = 3k+2 \end{cases}$$

bekommen.

## 5.2 Konvergenz in $\mathbb{R}$

**Lemma 5.2.1.** *Es seien $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ konvergente Folgen in $\mathbb{R}$ mit $x = \lim_{n\to\infty} x_n$ und $y = \lim_{n\to\infty} y_n$, dann sind auch $(x_n \pm y_n)_{n\in\mathbb{N}}$ und $(x_n y_n)_{n\in\mathbb{N}}$ konvergent und*

$$x \pm y = \lim_{n\to\infty} (x_n \pm y_n)_{n\in\mathbb{N}} \quad \text{und} \quad xy = \lim_{n\to\infty} (x_n y_n)_{n\in\mathbb{N}}.$$

*Gilt zusätzlich $y \neq 0$, dann gibt es ein $K \in \mathbb{N}$ mit $y_n \neq 0$ für alle $n \geq K$ und dann ist auch $\left(\frac{x_n}{y_n}\right)_{\substack{n\in\mathbb{N}\\ n\geq K}}$ konvergent mit $\frac{x}{y} = \lim_{n\to\infty} \frac{x_n}{y_n}$.*

*Beweis.* Ist $x = \lim_{n\to\infty} x_n$ und $y = \lim_{n\to\infty} y_n$ und $\varepsilon > 0$, so gibt es $k = k(\varepsilon) \in \mathbb{N}$ mit $|x - x_n| < \frac{\varepsilon}{2}$ für alle $n \geq k$ und es gibt $l = l(\varepsilon) \in \mathbb{N}$ mit $|y - y_n| < \frac{\varepsilon}{2}$ für $n \geq l$. Nach Dreiecksungleichung folgt

$$|x \pm y - (x_n \pm y_n)| \leq |x - x_n| + |y - y_n| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad \text{für } n \geq \max\{k, l\},$$

also $x \pm y = \lim_{n\to\infty} (x_n \pm y_n)$.

Nach Lemma 5.1.8 ist $C := \sup_{n\in\mathbb{N}} |y_n| < \infty$. Zu $\varepsilon > 0$ ist dann $\tilde{\varepsilon} := \dfrac{\varepsilon}{2\max\{|x|, C\} + 1} > 0$, daher gibt es $k \in \mathbb{N}$ mit $|x - x_n|, |y - y_n| < \tilde{\varepsilon}$ für $n \geq k$, damit ist

$$|xy - x_n y_n| = |x(y - y_n) + (x - x_n)y_n| \leq |x||y - y_n| + |x - x_n||y_n| \leq |x|\tilde{\varepsilon} + C\tilde{\varepsilon} \leq \varepsilon$$

für $n \geq k$, also $xy = \lim_{n\to\infty} x_n y_n$.

Ist $0 \neq y = \lim_{n\to\infty} y_n$, dann ist $0 \notin ]y - \frac{|y|}{2}, y + \frac{|y|}{2}[ = \{z \in \mathbb{R} : |z - y| < \frac{|y|}{2}\}$ und daher existiert ein $K \in \mathbb{N}$ mit $y_n \in ]y - \frac{|y|}{2}, y + \frac{|y|}{2}[$ für alle $n \geq K$; insbesondere ist $|y_n| > \frac{|y|}{2}$ für $n \geq K$. Wegen $x = \lim_{n\to\infty} x_n$ und $y = \lim_{n\to\infty} y_n$ gibt es für alle $\varepsilon > 0$ ein $k(\varepsilon) \in \mathbb{N}$ mit $|y - y_n| < \dfrac{\varepsilon|y|^2}{4|x| + 1}$ und $|x - x_n| < \dfrac{\varepsilon|y|}{4}$ für $n \geq k(\varepsilon)$; damit ist

$$\begin{aligned}\left|\frac{x}{y} - \frac{x_n}{y_n}\right| &= \left|\frac{xy_n - x_n y}{yy_n}\right| = \left|\frac{x(y_n - y) + y(x - x_n)}{yy_n}\right| \leq \frac{|x||y_n - y| + |y||x - x_n|}{|y||y_n|} \\ &\leq \frac{2|x|}{|y|^2}|y_n - y| + \frac{2}{|y|}|x - x_n| \leq \varepsilon\end{aligned}$$

für $n \geq \max\{K, k(\varepsilon)\}$ und damit $\frac{x}{y} = \lim_{n\to\infty} \frac{x_n}{y_n}$. □

**Lemma 5.2.2.** *Es seien $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ konvergente Folgen in $\mathbb{R}$ mit $x = \lim_{n\to\infty} x_n$, $y = \lim_{n\to\infty} y_n$ und $x_n \leq y_n$ für alle $n \in \mathbb{N}, n \geq N$. Dann gilt $x \leq y$.*

*Beweis.* Für jedes $\varepsilon > 0$ gibt es $M(\varepsilon) \in \mathbb{N}$ mit $|x - x_n| < \varepsilon$ für alle $n \geq M(\varepsilon)$ und es gibt $N(\varepsilon) \in \mathbb{N}$ mit $|y - y_n| < \varepsilon$ für $n \geq N(\varepsilon)$. Wegen $x_n \leq y_n$ für $n \geq N$ folgt $x - \varepsilon < x_n \leq y_n < y + \varepsilon$ für alle $n \geq \max\{N, M(\varepsilon), N(\varepsilon)\}$. Damit hat man für alle $\varepsilon > 0$ die Abschätzung $x - y < 2\varepsilon$ gezeigt, was nur für $x \leq y$ möglich ist. □

**Lemma 5.2.3.** *Es seien $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ konvergente Folgen in $\mathbb{R}$ mit $x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n$ und $x_n \le y_n$ für jedes $n \in \mathbb{N}$. Dann ist jede Folge $(z_n)_{n\in\mathbb{N}}$ mit $x_n \le z_n \le y_n$ für alle $n \in \mathbb{N}$ konvergent und $x = \lim_{n\to\infty} z_n$.*

*Beweis.* Wegen $x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n$ gibt es für alle $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$, so daß $|x - x_n| < \varepsilon$ und $|x - y_n| < \varepsilon$ für alle $n \ge N(\varepsilon)$ gilt. Aus $x_n \le z_n \le y_n$ folgt also

$$|x - z_n| \le \left\{ \begin{array}{ll} |x - y_n| & \text{für } x < x_n \\ \max\{|x - x_n|, |x - y_n|\} & \text{für } x \in [x_n, y_n] \\ |x - x_n| & \text{für } x > y_n \end{array} \right\} \le \varepsilon$$

für alle $n \ge N(\varepsilon)$, also ist $x = \lim_{n\to\infty} z_n$. □

**Lemma 5.2.4.** *Ist $(x_n)_{n\in\mathbb{N}}$ eine monoton steigende, nach oben beschränkte Folge in $\mathbb{R}$, dann ist $(x_n)_{n\in\mathbb{N}}$ konvergent mit*

$$\sup\{x_n : n \in \mathbb{N}\} = \lim_{n\to\infty} x_n \tag{5.2.1}$$

*Beweis.* $(x_n)_{n\in\mathbb{N}}$ ist genau dann nach oben beschränkt, wenn $X := \{x_n : n \in \mathbb{N}\}$ nach oben beschränkt ist. Nach Satz 4.1.4 existiert dann $\gamma := \sup(X) \in \mathbb{R}$ und nach Lemma 4.1.5 gibt es für jedes $k \in \mathbb{N}$ ein $x_{n_k} \in X$ mit $\gamma - \dfrac{1}{k} < x_{n_k} \le \gamma$. Da $(x_n)_{n\in\mathbb{N}}$ monoton steigend ist, gilt $\gamma - \frac{1}{k} < x_{n_k} \le x_m \le \gamma$ für alle $m \in \mathbb{N}$ mit $m \ge n_k$. Wegen Lemma 4.1.2 gibt es zu jedem $\varepsilon > 0$ ein $k \in \mathbb{N}$ mit $\dfrac{1}{k} < \varepsilon$, damit ist $|\gamma - x_m| < \dfrac{1}{k} < \varepsilon$ für alle $m \ge n_k$ oder $\sup\{x_n : n \in \mathbb{N}\} = \lim_{n\to\infty} x_n$. □

**Korollar 5.2.5.** *Ist $(x_n)_{n\in\mathbb{N}}$ eine monoton fallende, nach unten beschränkte Folge in $\mathbb{R}$, dann ist $(x_n)_{n\in\mathbb{N}}$ konvergent mit*

$$\inf\{x_n : n \in \mathbb{N}\} = \lim_{n\to\infty} x_n \tag{5.2.2}$$

**Beispiel 5.2.6.** Ist $|y| > 1$, so gilt $0 = \lim_{n\to\infty} \dfrac{n^k}{y^n}$ für alle $k \in \mathbb{N}$.

*Beweis.* Schreibe $|y| = 1 + x$ mit $x > 0$, dann gilt nach der Binomialformel

$$|y|^n = (1+x)^n = \sum_{l=0}^{n} \binom{n}{l} x^l > \binom{n}{k+1} x^{k+1} = \frac{n(n-1)\cdots(n-k)}{(k+1)!} x^{k+1}$$

für jedes $n \ge k+1$. Setzt man noch $n > 2k$ voraus, so ist $n - k > \dfrac{n}{2}$ und damit

$$\frac{n(n-1)\cdots(n-k)}{(k+1)!} x^{k+1} > \frac{n^{k+1} x^{k+1}}{2^{k+1}(k+1)!}$$

oder

$$0 \le \left| \frac{n^k}{y^n} \right| < \frac{2^{k+1}(k+1)!}{x^{k+1}} \cdot \frac{1}{n}$$

für $n > 2k$, woraus $0 = \lim_{n\to\infty} \dfrac{n^k}{y^n}$ folgt. □

**Beispiel 5.2.7.** Es sei $b \in \mathbb{N}$ mit $b \geq 2$ und $x \in ]0,1[$. Definiere eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\{0,1,...,b-1\}$ rekursiv durch

$$\begin{aligned} x_1 &:= \max\left\{n_1 \in \{0,1,...,b-1\} : \frac{n_1}{b} \leq x\right\} \\ &\vdots \\ x_{k+1} &:= \max\left\{n_{k+1} \in \{0,1,...,b-1\} : \frac{x_1}{b} + \frac{x_2}{b^2} + ... + \frac{x_{k-1}}{b^{k-1}} + \frac{x_k}{b^k} + \frac{n_{k+1}}{b^{k+1}} \leq x\right\} \end{aligned}$$

dann ist die Folge $\left(\sum_{k=1}^n \frac{x_k}{b^k}\right)_{n\in\mathbb{N}}$ monoton steigend und durch $x$ nach oben beschränkt, also konvergent. Es gilt sogar $x = \lim_{n\to\infty} \sum_{k=1}^n \frac{x_k}{b^k}$, denn nach Definition von $x_n$ ist

$$\sum_{k=1}^n \frac{x_k}{b^k} \leq x < \sum_{k=1}^{n-1} \frac{x_k}{b^k} + \frac{x_n+1}{b^n}$$

also

$$0 \leq x - \sum_{k=1}^n \frac{x_k}{b^k} = \left| x - \sum_{k=1}^n \frac{x_k}{b^k} \right| < \sum_{k=1}^{n-1} \frac{x_k}{b^k} + \frac{x_n+1}{b^n} - \sum_{k=1}^n \frac{x_k}{b^k} = \frac{1}{b^n}$$

Nach Lemma 4.1.2 ist $0 = \lim_{n\to\infty} \frac{1}{b^n}$ und damit $x = \lim_{n\to\infty} \sum_{k=1}^n \frac{x_k}{b^k}$ nach Lemma 5.2.3.

**Bemerkung 5.2.8.**

a) Für die $b$-adische Entwicklung $\left(\sum_{k=1}^n \frac{x_k}{b^k}\right)_{n\in\mathbb{N}}$, die ja durch die Folge der Zahlen $x_1, x_2, x_3, ... \in \{0,1,...,b-1\}$ gegeben ist schreibt man stattdessen $0,x_1x_2x_3...$ und kennzeichnet $0,x_1...x_{N-1}\overline{x_N...x_M}$, wenn sich in der Entwicklung die Zahlen $x_N...x_M$ in dieser Reihenfolge immer wieder wiederholen. Für $b = 10$ hat man die gewohnte Dezimalbruchentwicklung. In der Dezimalbruchentwicklung hat man etwa $\frac{1}{3} = 0,\overline{3}$. Nach obigen Konventionen steht $0,\overline{3}$ für den Grenzwert $\lim_{n\to\infty} \sum_{k=1}^n \frac{3}{10^k}$. Nach Beispiel 2.1.4 ist $\sum_{k=1}^n \frac{1}{10^k} = \frac{1-\frac{1}{10^{n+1}}}{1-\frac{1}{10}} - 1 \overset{n\to\infty}{\longrightarrow} \frac{1}{1-\frac{1}{10}} - 1 = \frac{1}{9}$.

b) Es sei

$$\begin{aligned} \mathcal{B} \;:=\; &\{0,x_1x_2x_3... : x_k \in \{0,1,...,b-1\} \text{ für alle } k \in \mathbb{N} \text{ und es gibt kein } N \in \mathbb{N} \text{ mit} \\ &\qquad x_k = b-1 \text{ für alle } k \geq N\} \end{aligned}$$

dann ist

$$\begin{aligned} f : \mathcal{B} &\to [0,1[ \\ 0,x_1x_2x_3... &\mapsto x := \lim_{n\to\infty} \sum_{k=1}^n \frac{x_k}{b^k} \end{aligned}$$

bijektiv.

*Beweis.* Es sei $x \in [0, 1[$ und $(x_k)_{k\in\mathbb{N}}, (y_k)_{k\in\mathbb{N}}$ Folgen in $\{0, 1, ..., b-1\}$ mit

$$x = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{x_k}{b^k} = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{y_k}{b^k}$$

und $\{k \in \mathbb{N} : x_k \neq y_k\} \neq \emptyset$. Es sei $K := \min\{k \in \mathbb{N} : x_k \neq y_k\}$ und oE. sei $x_K > y_K$. Für

$$\xi := \sum_{k=1}^{K} \frac{y_k}{b^k}$$

ist dann

$$0 \leq x - \xi = \lim_{n\to\infty} \sum_{k=K+1}^{n} \frac{y_k}{b^k} = \frac{x_K - y_K}{b^K} + \lim_{n\to\infty} \sum_{k=K+1}^{n} \frac{x_k}{b^k}. \tag{5.2.3}$$

Aus (5.2.3) folgt

$$x - \xi \leq \lim_{n\to\infty} \sum_{k=K+1}^{n} \frac{b-1}{b^k} = \frac{b-1}{b^{K+1}} \lim_{n\to\infty} \sum_{k=0}^{n-K-1} \frac{1}{b^k} = \frac{b-1}{b^{K+1}} \cdot \frac{1}{1-\frac{1}{b}} = \frac{1}{b^K} \tag{5.2.4}$$

und $x - \xi = \dfrac{1}{b^K}$ genau dann, wenn $b - 1 = y_k$ für alle $k \geq K + 1$ gilt. Aus (5.2.3) folgt wegen $x_K > y_K$ ebenso

$$x - \xi = \frac{x_K - y_K}{b^K} + \lim_{n\to\infty} \sum_{k=K+1}^{n} \frac{x_k}{b^k} \geq \frac{x_K - y_K}{b^K} \geq \frac{1}{b^K} \tag{5.2.5}$$

und $x - \xi = \dfrac{1}{b^K}$ gilt genau dann, wenn $x_K = y_K + 1$ und $x_k = 0$ für alle $k \geq K + 1$ gilt. Die beiden Abschätzungen (5.2.4) und (5.2.5) lassen sich nur für $x - \xi = \dfrac{1}{b^K}$ erfüllen. Wegen $K := \min\{k \in \mathbb{N} : x_k \neq y_k\}$ gilt $x_l = y_l$ für $l = 1, ..., K-1$ und $x_l = 0$ für $l \geq K + 1$, $y_l = b - 1$ für $l \geq K + 1$. Von den beiden Möglichkeiten

$$x = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{x_k}{b^k} = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{y_k}{b^k}$$

als $b-$adischen Bruch zu schreiben ist $(y_n)_{n\in\mathbb{N}} \notin \mathcal{B}$ und damit $f$ injektiv. Das Verfahren aus Beispiel 5.2.7 definiert zu $x \in [0, 1[$ eine Folge $(x_n)_{n\in\mathbb{N}} \in \mathcal{B}$ mit $x = \lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \dfrac{x_k}{b^k}$, weshalb $f$ surjektiv ist. □

c) Unter Verwendung von Teil b) sieht man: Ist $a, b \in \mathbb{R}$ mit $a < b$, dann ist das Intervall $]a, b[$ überabzählbar.

*Beweis.* Der Beweis von (c) bleibt als Aufgabe. □

## 5.3 Wurzeln

**Beispiel 5.3.1.** Die Gleichung $X^2 = 2$ hat in $\mathbb{Q}$ keine Lösung.

*Beweis.* Angenommen es gibt eine rationale Zahl $q = \frac{n}{m}$ mit $q^2 = 2$. Nach Kürzen aller Potenzen von 2 dürfen wir oE. annehmen, daß mindestens eine der Zahlen $n, m$ ungerade ist. Es gilt dann $n^2 = 2m^2$, also ist $n^2$ gerade und damit auch $n$ gerade. Dann ist aber $n^2$ durch 4 teilbar, also ist auch $2m^2$ durch 4 teilbar. Das ist aber nur möglich, wenn $m^2$ und damit $m$ gerade ist. Dann sind aber $m$ und $n$ gerade, im Widerspruch dazu, daß alle gemeinsamen Faktoren gekürzt sind. □

**Satz 5.3.2.** *Ist $n \in \mathbb{N}$ und $x \in \mathbb{R}$ mit $x > 0$, dann gibt es genau ein $\omega > 0$ mit $\omega^n = x$.*

*Beweis.* Eindeutigkeit: Sind $\omega_1, \omega_2 \in \mathbb{R}$, $\omega_1, \omega_2 > 0$ mit $\omega_1^n = \omega_2^n = x$ und $\omega_1 \neq \omega_2$, dann dürfen wir oE. $\omega_1 < \omega_2$ annehmen. Da $\begin{array}{rcl} y^n : \mathbb{R}_+ & \to & \mathbb{R}_+ \\ y & \mapsto & y^n \end{array}$ streng monoton steigend ist, folgt dann $x = \omega_1^n < \omega_2^n = x$; Widerspruch.
Existenz:

$$\emptyset \neq X := \{y \in \mathbb{R}_+ : x \leq y^n\} \subseteq \mathbb{R}$$

ist nach unten beschränkt, etwa 0 ist eine untere Schranke und für $x \geq 1$ ist $x \in X \neq \emptyset$ und für $x < 1$ ist $1 \in X \neq \emptyset$. Daher existiert $\omega := \inf(X)$ in $\mathbb{R}$.

- Es gilt $\omega > 0$, denn wegen $\omega = \inf\{y \in [0, \infty[ : x \leq y^n\} \geq 0$ bleibt sonst nur $\omega = 0$, dann gibt es nach Korollar 4.1.7 für jedes $k \in \mathbb{N}$ ein $y_k \in X$ mit $0 \leq y_k < \frac{1}{k}$, weshalb $x \leq y_k^n < \frac{1}{k^n} \leq \frac{1}{k}$ nach Lemma 3.4.3 gilt. Damit ist $x \in \bigcap_{k \in \mathbb{N}} \left[0, \frac{1}{k}\right[ = \{0\}$, was im Widerspruch zu $x > 0$ steht.

- Für $y \in X$ und $z \geq y$ ist $x \leq y^n \leq z^n$, dh. $z \in X$. Nach Charakterisierung des Infimums in Korollar 4.1.7 gibt es für jedes $k \in \mathbb{N}$ ein $x_k \in X$ mit $\inf(X) = \omega \leq x_k < \omega + \frac{1}{k}$. Somit ist für jedes $k \in \mathbb{N}$ auch
$$\omega + \frac{1}{k} \in X. \tag{5.3.1}$$

- $\omega \in X$:
Für alle $m \in \mathbb{N}$ ist
$$\left(\omega + \frac{1}{m}\right)^n = \omega^n + \sum_{l=1}^{n} \binom{n}{l} \frac{1}{m^l} \omega^{n-l} \leq \omega^n + \frac{1}{m} \sum_{l=1}^{n} \binom{n}{l} \omega^{n-l}.$$
Angenommen $\omega^n < x$, dann gibt es wegen $\sum_{l=1}^{n} \binom{n}{l} \omega^{n-l} > 0$ nach Lemma 4.1.2 ein $k \in \mathbb{N}$ mit $\frac{1}{k} \sum_{l=1}^{n} \binom{n}{l} \omega^{n-l} < x - \omega^n$, also ist $\left(\omega + \frac{1}{k}\right)^n < x$, dh. $\omega + \frac{1}{k} \notin X$ im Widerspruch zu (5.3.1).

- $x \geq \omega^n$,
  denn sonst ist $\dfrac{x}{\omega^n} < 1$ und damit gibt es nach Lemma 4.1.2 ein $k \in \mathbb{N}$ mit $\dfrac{n}{\omega k} < 1 - \dfrac{x}{\omega^n}$, also ist $0 \leq \dfrac{x}{\omega^n} < 1 - \dfrac{n}{\omega k} < 1$. Da also
  $$\frac{1}{\omega k} \leq \frac{n}{\omega k} \leq 1 \tag{5.3.2}$$
  ist, dürfen wir die Bernoulli-Ungleichung anwenden und bekommen:
  $$x < \omega^n \left(1 - \frac{n}{\omega k}\right) \leq \omega^n \left(1 - \frac{1}{\omega k}\right)^n = \left(\omega - \frac{1}{k}\right)^n.$$
  Damit ist wegen (5.3.2) $\omega - \frac{1}{k} > 0$, also auch $\omega - \dfrac{1}{k} \in X$ im Widerspruch zu $\omega = \inf(X)$.

Aus $\omega \in X$, dh. $x \leq \omega^n$ und $x \geq \omega^n$ folgt nun $x = \omega^n$. □

**Definition 5.3.3.** *Ist $x \in \mathbb{R}$ mit $x > 0$ und $n \in \mathbb{N}$, so heißt die Zahl $\omega > 0$ mit $\omega^n = x$ aus Satz 5.3.2 die $n$-te Wurzel aus $x$. Schreibweise : $\sqrt[n]{x}$ oder $x^{\frac{1}{n}}$.*

**Lemma 5.3.4.** *Sind $x, y \in \mathbb{R}$, $x > 0$ und $y > 0$, $m, n \in \mathbb{N}$, dann gilt:*

*a)* $\sqrt[n]{xy} = \sqrt[n]{x}\,\sqrt[n]{y}$

*b) Ist $x \leq y$, so gilt $\sqrt[n]{x} \leq \sqrt[n]{y}$.*

*c)* $\sqrt[m]{\sqrt[n]{x}} = \sqrt[mn]{x}$.

*Beweis.* Sind $x, y \in \mathbb{R}$, $x > 0$ und $y > 0$, $m, n \in \mathbb{N}$, dann gilt:

a) Nach Satz 5.3.2 ist $\sqrt[n]{x} > 0$ und $\sqrt[n]{y} > 0$, ferner $(\sqrt[n]{x}\,\sqrt[n]{y})^n = (\sqrt[n]{x})^n(\sqrt[n]{y})^n = xy$, was zeigt, daß $\sqrt[n]{x}\,\sqrt[n]{y}$ die definierenden Eigenschaften von $\sqrt[n]{xy}$ erfüllt.

b) Angenommen $x \leq y$ und $\sqrt[n]{x} > \sqrt[n]{y}$, dann ist nach Lemma 3.4.3 auch $x = (\sqrt[n]{x})^n > (\sqrt[n]{y})^n = y$ im Widerspruch zur Voraussetzung $x \leq y$.

c) Nach Definition ist $\sqrt[m]{\sqrt[n]{x}} > 0$ und $(\sqrt[m]{\sqrt[n]{x}})^{mn} = ((\sqrt[m]{\sqrt[n]{x}})^m)^n = (\sqrt[n]{x})^n = x$, also erfüllt $\sqrt[m]{\sqrt[n]{x}}$ wieder alle definierenden Eigenschaften von $\sqrt[mn]{x}$. □

**Satz 5.3.5.** *a) Für alle $p > 0$ gilt: $\lim\limits_{n\to\infty} \sqrt[n]{p} = 1$.*

*b)* $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$

*Beweis.* a) Ist $p \geq 1$, so ist auch $\sqrt[n]{p} \geq 1$, also $a_n := \sqrt[n]{p} - 1 \geq 0$. Nach der Bernoulli-Ungleichung gilt also $p = (1 + a_n)^n \geq 1 + na_n > na_n$ oder $0 \leq a_n \leq \frac{p}{n}$ für alle $n \in \mathbb{N}$, also $\lim\limits_{n\to\infty} a_n = 0$.

Ist $0 < p < 1$, so ist $\sqrt[n]{p} < 1$, also $a_n := \frac{1}{\sqrt[n]{p}} - 1 > 0$ und $\sqrt[n]{p} = \frac{1}{1+a_n}$. Mit der Bernoulli-Ungleichung folgt nun wieder: $(1 + a_n)^n \geq 1 + na_n > na_n$, also ist $p = \frac{1}{(1+a_n)^n} < \frac{1}{na_n}$ und damit $0 \leq a_n \leq \frac{1}{pn}$, woraus wieder $\lim\limits_{n\to\infty} a_n = 0$ folgt.

b) Für $n \geq 2$ ist $b_n := \sqrt[n]{n} - 1 > 0$ damit folgt:

$$n = (\sqrt[n]{n})^n = (b_n + 1)^n = \sum_{j=0}^{n} \binom{n}{j} b_n^j > \binom{n}{2} b_n^2 = \frac{n(n-1)}{2} b_n^2,$$

denn $\binom{n}{2} b_n^2$ ist ein Summand in dere Summe mit lauter positiven Termen. Damit ist $|b_n - 0| = b_n < \sqrt{\frac{2}{n-1}}$ für $n \geq 2$. Für jedes $\varepsilon > 0$ und $n \geq N = N(\varepsilon) := \max\{2, \frac{2}{\varepsilon^2} + 1\}$ gilt: $\frac{2}{n-1} \leq \frac{2}{N-1} \leq \varepsilon^2$, d.h. $|b_n| < \sqrt{\frac{2}{n-1}} \leq \varepsilon$ für $n \geq N(\varepsilon)$ und damit $\lim_{n\to\infty} b_n = 0$ oder $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

□

**Lemma 5.3.6.** *Für $p > 0$ und $a_0 > 0$ definiere man rekursiv die Folge $(a_n)_{n\in\mathbb{N}_0}$ durch*

$$a_{n+1} = \frac{1}{2}(a_n + \frac{p}{a_n}).$$

*Dann konvergiert die Folge $(a_n)_{n\in\mathbb{N}}$ gegen $\sqrt{p}$.*

*Beweis.* Wir bemerken zunächst $a_n > 0$, denn nach Voraussetzung ist $p, a_0 > 0$ und damit $a_{n+1} = \frac{1}{2}(a_n + \frac{p}{a_n}) > 0$ für $n \in \mathbb{N}_0$. Für $n \geq 1$ ist

$$a_n^2 - p = \frac{1}{4}(a_{n-1} + \frac{p}{a_{n-1}})^2 - p = \frac{1}{4}(a_{n-1} - \frac{p}{a_{n-1}})^2 \geq 0$$

oder $a_n \geq \sqrt{p}$. Ferner gilt

$$a_n - a_{n+1} = a_n - \frac{1}{2}(a_n + \frac{p}{a_n}) = \frac{1}{2}(a_n - \frac{p}{a_n}) = \frac{1}{2a_n}(a_n^2 - p) \geq 0,$$

also ist $(a_n)_{n\in\mathbb{N}}$ eine monoton fallende, nach unten beschränkte Folge mit einer unteren Schranke $\sqrt{p}$. Folglich existiert der Limes $a := \lim_{n\to\infty} a_n = \inf\{a_n : n \in \mathbb{N}\}$ und es gilt $a \geq \sqrt{p}$. Nach Definition des Grenzwerts [2] gilt dann auch $a = \lim_{n\to\infty} a_{n+1}$ und wegen $a \neq 0$ folgt aus den Rechenregeln für Summe und Quotient:

$$a = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \frac{1}{2}(a_n + \frac{p}{a_n}) = \frac{a}{2} + \frac{p}{2a}$$

dh. der Grenzwert $a$ erfüllt die Gleichung $a^2 = p$. Wegen $a \geq 0$ bedeutet dies aufgrund der Definition der Wurzel $a = \sqrt{p}$. □

**Bemerkung 5.3.7.** Wählen wir in Lemma 5.3.6 $p = 2$ und $a_0 = 1$, so gilt offenbar $a_n \in \mathbb{Q}$ für alle $n \in \mathbb{N}$. Damit ist $(a_n)_{n\in\mathbb{N}}$ ein Beispiel für eine Cauchyfolge in $\mathbb{Q}$, die in $\mathbb{Q}$ nicht konvergiert, da $\sqrt{2} \notin \mathbb{Q}$ nach Beispiel 5.3.1.

---

[2] Denn existiert $a = \lim_{n\to\infty} a_n$, so bedeutet dies per Definition: Für jedes $\varepsilon > 0$ gibt es ein $N(\varepsilon) \in \mathbb{N}$, so daß $|a - a_n| < \varepsilon$ für jedes $n \in \mathbb{N}, n \geq N(\varepsilon)$. Für die Folge $(b_n)_{n\in\mathbb{N}}$ mit $b_n := a_{n+1}$ gilt dann auch $|a - b_n| < \varepsilon$ für alle $n \geq N(\varepsilon) + 1$, dh. $a = \lim_{n\to\infty} b_n = \lim_{n\to\infty} a_{n+1}$

## 5.4 Konvergenz in $\mathbb{C}$ und in normierten Räumen

**Definition 5.4.1.** *Zu $z \in \mathbb{C}$ heißt $|z| := \sqrt{z\bar{z}}$ der* **Absolutbetrag** *von $z$.*

**Lemma 5.4.2.** *Für den Absolutbetrag $|\cdot| : \mathbb{C} \to \mathbb{R}$ gilt:*

*a)* $|z| \geq 0$ *für alle* $z \in \mathbb{C}$.

*b)* $|z| = 0$ *genau dann wenn* $z = 0 + i0$ *ist.*

*c) Für alle* $z \in \mathbb{C}$ *ist* $|\mathrm{Re}\, z| \leq |z|$ *und* $|\mathrm{Im}\, z| \leq |z|$.

*d) Für alle* $w, z \in \mathbb{C}$ *gilt* $|wz| = |w| \cdot |z|$.

*e) Für alle* $w, z \in \mathbb{C}$ *gilt die* **Dreiecksungleichung**: $|w + z| \leq |w| + |z|$.

*Beweis.* Nach Lemma 4.2.5 sind a) und b) offensichtlich. Für $z = x + iy \in \mathbb{C}$ und $w \in \mathbb{C}$ gilt:

c) $|\mathrm{Re}\, z| = |x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2} = |z|$ und $|\mathrm{Im}\, z| = |y| = \sqrt{y^2} \leq \sqrt{x^2 + y^2} = |z|$ wegen Monotonie der Wurzel.

d) Nach Definition ist $|wz|^2 = (wz)(\overline{wz}) = (z\bar{z})(w\bar{w}) = |z|^2|w|^2$, also $|wz| = |w| \cdot |z|$.

e) Folgt durch Bilden der Quadratwurzel aus

$$\begin{aligned} |w + z|^2 &= (z + w)(\overline{w + z}) = z\bar{z} + z\bar{w} + w\bar{z} + w\bar{w} = z\bar{z} + 2\mathrm{Re}\,(z\bar{w}) + w\bar{w} \\ &\leq |z|^2 + 2|z\bar{w}| + |w|^2 = |z|^2 + 2|z||w| + |w|^2 = (|z| + |w|)^2. \end{aligned}$$

□

**Definition 5.4.3.** *Es sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und $V$ ein $\mathbb{K}$-Vektorraum. Eine Abbildung* $\begin{array}{rcl} \|\cdot\| : V & \to & [0, \infty[ \\ v & \mapsto & \|v\| \end{array}$ *heißt* **Norm**, *wenn*

*a) Aus* $\|v\| = 0$ *folgt* $v = \mathbf{0}$

*b) Für alle* $\lambda \in \mathbb{K}$ *und* $v \in V$ *gilt:*

$$\|\lambda v\| = |\lambda| \cdot \|v\| \tag{5.4.1}$$

*c) Für alle* $v, w \in V$ *gilt die* **Dreiecksungleichung:**

$$\|v + w\| \leq \|v\| + \|w\| \tag{5.4.2}$$

*Ist $\|\cdot\|$ eine Norm auf dem $\mathbb{K}$-Vektorraum $V$, so heißt $(V, \|\cdot\|)$ ein* **normierter Vektorraum**.

**Korollar 5.4.4.** $\begin{array}{rcl} |\cdot| : \mathbb{C} & \to & [0, \infty[ \\ z & \mapsto & |z| \end{array}$ *ist eine Norm auf* $\mathbb{C}$.

**Beispiel 5.4.5.** Ist $d \in \mathbb{N}$, so werden auf $\mathbb{K}^d$ durch

$$\left\| \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \right\|_1 := |x_1| + ... + |x_n| \tag{5.4.3}$$

und

$$\left\| \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \right\|_\infty := \max\{|x_1|, ..., |x_d|\} \tag{5.4.4}$$

zwei Normen $\begin{array}{rcl} \|\cdot\|_1 : \mathbb{K}^d & \to & [0,\infty[ \\ \underline{x} & \mapsto & \|\underline{x}\|_1 \end{array}$ und $\begin{array}{rcl} \|\cdot\|_\infty : \mathbb{K}^d & \to & [0,\infty[ \\ \underline{x} & \mapsto & \|\underline{x}\|_\infty \end{array}$ definiert.

**Lemma 5.4.6.** *Ist $(V, \|\cdot\|)$ ein normierter Vektorraum, so definiert*

$$\begin{array}{rcl} d_{\|\cdot\|} : V \times V & \to & [0,\infty[ \\ (v,w) & \mapsto & \|v-w\| \end{array}$$

*eine Metrik auf $V$. $d_{\|\cdot\|}$ heißt die* **von der Norm $\|\cdot\|$ induzierte Metrik**.

*Beweis.* Nach den Eigenschaften der Norm gilt für alle $u, v, w \in V$:

- $d_{\|\cdot\|}(u,v) = \|u-v\| \geq 0$ und $d_{\|\cdot\|}(u,v) = \|u-v\| = 0$ genau dann wenn $u - v = \mathbf{0}$, dh. $u = v$ ist.
- $d_{\|\cdot\|}(u,v) = \|u-v\| = \|-1(v-u)\| = \|v-u\| = d_{\|\cdot\|}(v,u)$.
- $d_{\|\cdot\|}(u,w) = \|u-w\| = \|(u-v)+(v-w)\| \leq \|u-v\|+\|v-w\| = d_{\|\cdot\|}(u,v)+d_{\|\cdot\|}(v,w)$.

□

**Definition 5.4.7.** *Ist $(V, \|\cdot\|)$ ein normierter $\mathbb{K}-$ Vektorraum, dann heißt*

- *eine Folge $(x_n)_{n\in\mathbb{N}}$ in $V$* **konvergent**, *wenn sie bezüglich der von der Norm $\|\cdot\|$ induzierten Metrik $d_{\|\cdot\|}$ konvergiert; das bedeutet hier also, wenn es ein $x \in V$ gibt, so daß es für jedes $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$ gibt mit $\|x - x_n\| < \varepsilon$ für alle $n \geq N(\varepsilon)$.*
- *eine Folge $(x_n)_{n\in\mathbb{N}}$ in $V$ eine* **Cauchyfolge**, *wenn sie bezüglich der von der Norm $\|\cdot\|$ induzierten Metrik $d_{\|\cdot\|}$ Cauchyfolge ist; das bedeutet hier also, daß es für jedes $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$ gibt mit $\|x_m - x_n\| < \varepsilon$ für alle $m, n \geq N(\varepsilon)$.*
- *$(V, \|\cdot\|)$ ein* **$\mathbb{K}$-Banachraum**, *wenn $(V, d_{\|\cdot\|})$ für die von der Norm $\|\cdot\|$ induzierte Metrik $d_{\|\cdot\|}$ vollständig ist.*

**Lemma 5.4.8.** *Eine Folge $(z_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ konvergiert genau dann, wenn die beiden reellen Folgen $(\operatorname{Re} z_n)_{n\in\mathbb{N}}$ und $(\operatorname{Im} z_n)_{n\in\mathbb{N}}$ konvergieren. In diesem Fall ist*

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty} \operatorname{Re} z_n + i \lim_{n\to\infty} \operatorname{Im} z_n.$$

*Eine Folge $(z_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ ist genau dann Cauchyfolge, wenn $(\operatorname{Re} z_n)_{n\in\mathbb{N}}$ und $(\operatorname{Im} z_n)_{n\in\mathbb{N}}$ Cauchyfolgen in $\mathbb{R}$ sind.*

*Beweis.* Ist $z = \lim_{n\to\infty} z_n$, dann gilt $0 = \lim_{n\to\infty} |z - z_n|$. Wegen der Abschätzungen $0 \leq |\operatorname{Re} z - \operatorname{Re} z_n| \leq |z - z_n|$ und $0 \leq |\operatorname{Im} z - \operatorname{Im} z_n| \leq |z - z_n|$ folgt $\operatorname{Re} z = \lim_{n\to\infty} \operatorname{Re} z_n$ und $\operatorname{Im} z = \lim_{n\to\infty} \operatorname{Im} z_n$ nach Lemma 5.2.3. Ist $x = \lim_{n\to\infty} \operatorname{Re} z_n$ und $y = \lim_{n\to\infty} \operatorname{Im} z_n$, dann ist

$$0 \leq |(x+iy) - z_n| = |[(x - \operatorname{Re} z_n) + i0] + [0 + i(y - \operatorname{Im} z_n)]| \leq |x - \operatorname{Re} z_n| + |i(y - \operatorname{Im} z_n)|,$$

nach Dreiecksungleichung und Definition des Absolutbetrags also $0 = \lim_{n\to\infty} |(x+iy) - z_n|$ nach Lemma 5.2.3. Die gleichen Abschätzungen plus $\varepsilon$-tik zeigen die Behauptung für die Cauchyfolgen. □

**Korollar 5.4.9.** $\mathbb{C}$ *ist vollständig.*

**Lemma 5.4.10.** *Es sei* $(\lambda_n)_{n\in\mathbb{N}}$ *eine konvergente Folge in* $\mathbb{K}$ *und* $\lambda := \lim_{n\to\infty} \lambda_n$. *Es sei* $(V, \|\cdot\|)$ *ein normierter* $\mathbb{K}$*-Vektorraum,* $(x_n)_{n\in\mathbb{N}}$ *und* $(y_n)_{n\in\mathbb{N}}$ *konvergente Folgen in* $V$ *mit* $x := \lim_{n\to\infty} x_n$ *und* $y := \lim_{n\to\infty} y_n$. *Dann konvergiert* $(x_n + \lambda_n y_n)_{n\in\mathbb{N}}$ *und es gilt:*

$$\lim_{n\to\infty} (x_n + \lambda_n y_n) = x + \lambda y. \tag{5.4.5}$$

*Beweis.* Da die Folge $(\lambda_n)_{n\in\mathbb{N}}$ konvergiert, ist $(|\lambda_n|)_{n\in\mathbb{N}}$ beschränkt, also ist

$$C := \sup\{|\lambda_n| : n \in \mathbb{N}\} < \infty.$$

Ist $\varepsilon > 0$, so gibt es da $\lambda := \lim_{n\to\infty} \lambda_n$, $x = \lim_{n\to\infty} x_n$ und $y = \lim_{n\to\infty} y_n$ ist ein $N = N(\varepsilon) \in \mathbb{N}$ mit $\|x_n - x\| < \frac{\varepsilon}{3}$, $\|y_n - y\| < \frac{\varepsilon}{3C+1}$ und $|\lambda_n - \lambda| < \frac{\varepsilon}{3(|y|+1)}$ für alle $n \geq N(\varepsilon)$. Nach den Eigenschaften der Norm folgt

$$\begin{aligned}
\|(x_n + \lambda_n y_n) - (x + \lambda y)\| &= \|(x_n - x) + \lambda_n(y_n - y) + (\lambda_n - \lambda)y\| \\
&\leq \|x_n - x\| + |\lambda_n| \cdot \|y_n - y\| + |\lambda_n - \lambda| \cdot \|y\| < \varepsilon
\end{aligned}$$

für alle $n \in \mathbb{N}$, $n \geq N(\varepsilon)$, dh. $\lim_{n\to\infty} (x_n + \lambda_n y_n) = x + \lambda y$. □

**Definition 5.4.11.** *Es sei* $V$ *ein* $\mathbb{K}$*-Vektorraum mit* $\mathbb{K} = \mathbb{R}$ *oder* $\mathbb{K} = \mathbb{C}$. *Eine Abbildung* $s : V \times V \to \mathbb{K}$ *heißt* **Sesquilinearform** *auf* $V$, *wenn für alle* $u, v, w \in V$ *und* $\lambda, \mu \in \mathbb{K}$ *gilt:*

$$s(u, \lambda v + \mu w) = \lambda s(u, v) + \mu s(u, w) \tag{5.4.1}$$

$$s(\lambda u + \mu v, w) = \overline{\lambda} s(u, w) + \overline{\mu} s(v, w). \tag{5.4.2}$$

*Eine Sesquilinearform* $s : V \times V \to \mathbb{K}$ *heißt*

- **hermitesch**, *wenn für alle* $v, w \in V$ *gilt:*

$$s(v, w) = \overline{s(w, v)}. \tag{5.4.3}$$

- **positiv definit**, *wenn* $s(v, v) \in \mathbb{R}$ *und* $s(v, v) \geq 0$ *für alle* $v \in V$ *gilt und darüberhinaus* $s(v, v) = 0$ *äquivalent zu* $v = \mathbf{0}$ *ist.*

*Ist* $\begin{array}{rcl} s: V\times V & \to & \mathbb{K} \\ (v,w) & \mapsto & s(v,w) \end{array}$ *eine Sesquilinearform, so heißt* $\begin{array}{rcl} q: V & \to & \mathbb{K} \\ v & \mapsto & s(v,v) \end{array}$ *die zu s gehörige* **quadratische Form**. *Eine hermitesche, positiv definite Sesquilinearform*

$$\langle\cdot,\cdot\rangle : V\times V \longrightarrow \mathbb{K}$$

*heißt ein* **Skalarprodukt** *in* $V$. *Ein* $\mathbb{R}$*-Vektorraum* $V$ *mit Skalarprodukt heißt* **euklidischer Vektorraum**, *ein* $\mathbb{C}$*-Vektorraum mit Skalarprodukt heißt* **unitärer Vektorraum**.

**Bemerkung 5.4.12.**

a) Im Fall $\mathbb{K}=\mathbb{R}$ ist eine Sesquilinearform (sesqui=eineinhalb) tatsächlich bilinear; aus (5.4.2) wird also $s(\lambda u+\mu v,w)=\lambda s(u,w)+\mu s(v,w)$.

b) Mit (5.4.1) und (5.4.2) folgen wir den Konventionen, wie sie in der mathematischen Physik gelten. In den meisten Mathematikbüchern wird stattdessen $s(u,\lambda v+\mu w)=\overline{\lambda}s(u,v)+\overline{\mu}s(u,w)$ und $s(\lambda u+\mu v,w)=\lambda s(u,w)+\mu s(v,w)$ gefordert.

**Lemma 5.4.13.** *Ist s eine Sesquilinearform auf einem* $\mathbb{K}$*-Vektorraum und q die zu s gehörige quadratische Form, dann gilt*

*a) Für alle* $v,w\in V$ *gilt die* **Parallelogrammidentität:**

$$q(v+w)+q(v-w)=2(q(v)+q(w)) \tag{5.4.4}$$

*b) Ist V ein* $\mathbb{C}$*-Vektorraum, so gilt für alle* $v,w\in V$ *die* **Polarisierungsidentität:**

$$s(v,w)=\frac{1}{4}(q(v+w)-q(v-w)+iq(v-iw)-iq(v+iw)) \tag{5.4.5}$$

*c) Ist s eine hermitesche Bilinearform auf einem* $\mathbb{R}$*-Vektorraum V, so gilt für alle* $v,w\in V$ *die Polarisierungsidentität:*

$$s(v,w)=\frac{1}{4}(q(v+w)-q(v-w)). \tag{5.4.6}$$

**Beispiel 5.4.14.**

a) Schreibe $\underline{x}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}, \underline{y}=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}\in\mathbb{C}^n$, dann definiert

$$\langle\underline{x},\underline{y}\rangle_{\mathbb{C}^n} := \overline{x_1}y_1+...+\overline{x_n}y_n$$

das **Standardskalarprodukt** im $\mathbb{C}^n$.

*Beweis.* Nach den elementaren Rechenregeln ist $\langle\cdot,\cdot\rangle$ offenbar eine hermitesche Sesquilinearform, deren quadratische Form wegen $\left\langle\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\right\rangle=\sum_{k=1}^n|x_k|^2$ stets nichtnegative Werte annimmt. Weil $\sum_{k=1}^n|x_k|^2=0$ zu $x_1=...=x_n=0$ äquivalent ist, ist $\langle\cdot,\cdot\rangle$ auch positiv definit. □

b) Für $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \underline{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ wird durch

$$\langle \underline{x}, \underline{y} \rangle_{\mathbb{R}^n} := x_1 y_1 + ... + x_n y_n$$

das **Standardskalarprodukt** im $\mathbb{R}^n$ definiert.

**Lemma 5.4.15.** *Es sei $V$ ein $\mathbb{K}$-Vektorraum mit Skalarprodukt und $u, v \in V$. Dann sind äquivalent:*

*a)* $u = v$

*b) Für alle $w \in V$ gilt: $\langle u, w \rangle = \langle v, w \rangle$.*

*Beweis.* Ist $u = v$, so ist $\langle u, w \rangle - \langle v, w \rangle = \langle u - v, w \rangle = \langle \mathbf{0}, w \rangle = 0$ für jedes $w \in V$. Ist umgekehrt $\langle u, w \rangle = \langle v, w \rangle$ für jedes $w \in V$, so ist $\langle u - v, w \rangle = 0$ für alle $w \in V$, also auch für $w = u - v$, also ist $\langle u - v, u - v \rangle = 0$ und damit $u - v = \mathbf{0}$. □

**Satz 5.4.16.** *Es sei $(V, \langle \cdot, \cdot \rangle)$ ein $\mathbb{K}$-Vektorraum mit Skalarprodukt und für $v \in V$ sei $\|v\| := \sqrt{\langle v, v \rangle}$, dann ist* $\begin{array}{rcl} \| \cdot \| : V & \to & [0, \infty[ \\ v & \mapsto & \|v\| \end{array}$ *eine Norm und dann gilt für alle $v, w \in V$ und $\lambda \in \mathbb{K}$:*

*a)* **Cauchy-Schwarzsche Ungleichung:**

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\| \tag{5.4.7}$$

*mit dem Zusatz: Die Gleichheit $|\langle v, w \rangle| = \|v\| \cdot \|w\|$ gilt genau dann, wenn $v = \mathbf{0}$ oder $w = \mathbf{0}$ oder wenn es $\lambda \in \mathbb{K}$ mit $v - \lambda w = \mathbf{0}$ gibt.*

*b)* **Satz von Pythagoras:**

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + \langle v, w \rangle + \langle w, v \rangle \tag{5.4.8}$$

*c)* **Parallelogrammgleichung:**

$$\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2. \tag{5.4.9}$$

*Beweis.* Für $w = 0$ gilt $|\langle v, w \rangle| = \|v\| \cdot \|w\| = 0$. Betrachte also nun den verbleibenden Fall $w \neq 0$. Aus der Definition des Skalarprodukt folgt für $\lambda \in \mathbb{K}$:

$$\begin{aligned} 0 &\leq \langle v - \lambda w, v - \lambda w \rangle = \langle v, v - \lambda w \rangle - \overline{\lambda} \langle w, v - \lambda w \rangle = \\ &= \langle v, v \rangle - \lambda \langle v, w \rangle - \overline{\lambda} \langle w, v \rangle + \overline{\lambda} \lambda \langle w, w \rangle = \\ &= \langle v, v \rangle - \lambda \langle v, w \rangle - \overline{\lambda} \overline{\langle v, w \rangle} + \overline{\lambda} \lambda \langle w, w \rangle \end{aligned} \tag{5.4.10}$$

Wegen $w \neq 0$ können wir hier $\lambda = \frac{\overline{\langle v, w \rangle}}{\langle w, w \rangle}$ wählen und für diese Wahl von $\lambda$ Abschätzung (5.4.10) mit $\langle w, w \rangle$ multiplizieren. Dann erhalten wir:

$$\begin{aligned} 0 &\leq \langle v, v \rangle \langle w, w \rangle - \langle v, w \rangle \overline{\langle v, w \rangle} - \langle v, w \rangle \overline{\langle v, w \rangle} + \langle v, w \rangle \overline{\langle v, w \rangle} = \\ &= \langle v, v \rangle \langle w, w \rangle - |\langle v, w \rangle|^2. \end{aligned}$$

Dies ist die Cauchy-Schwarzsche Ungleichung. In der ersten Zeile von (5.4.10) gilt $0 = \langle v - \lambda w, v - \lambda w\rangle$ genau dann, wenn $v - \lambda w = 0$ ist. Der Satz des Pythagoras folgt unmittelbar aus den Rechenregeln für das Skalarprodukt und die Parallelogrammgleichung ist ein Spezialfall von Lemma 5.4.13. Wegen der positiven Definitheit von $\langle\cdot,\cdot\rangle$ ist $\|v\| \geq 0$ für jedes $v \in V$ und $\|v\| = 0$ äquivalent zu $v = 0$. Für alle $\lambda \in \mathbb{K}$ und $v \in V$ gilt

$$\|\lambda v\|^2 = \langle \lambda v, \lambda v\rangle = \overline{\lambda}\lambda\langle v, v\rangle = |\lambda|^2\|v\|^2.$$

Sind $v, w \in V$, dann folgt

$$\begin{aligned} \|v + w\|^2 &= \|v\|^2 + \|w\|^2 + 2\mathrm{Re}\,\langle v, w\rangle \leq \|v\|^2 + \|w\|^2 + 2|\langle v, w\rangle| \\ &\leq \|v\|^2 + \|w\|^2 + 2\|v\| \cdot \|w\| = (\|v\| + \|w\|)^2 \end{aligned}$$

aus der Cauchy-Schwarzschen Ungleichung, woraus die Dreiecksungleichung folgt. □

**Bemerkung 5.4.17.** Die in Satz 5.4.16 definierte Norm heißt die **vom Skalarprodukt induzierte Norm** und wenn nichts anderes gesagt wird, betrachtet man jeden $\mathbb{K}$-Vektorraum mit Skalarprodukt als einen normierten Vektorraum bezüglich dieser Norm. Ein $\mathbb{K}-$ Vektorraum $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ mit Skalarprodukt, der bezüglich dieser induzierten Norm vollständig ist heißt ein **Hilbertraum** . Die Eigenschaften von Satz 5.4.16 zeichnen eine Norm, die von einem Skalarprodukt induziert wird gegenüber einer „gewöhnlichen“ Norm aus. Es gilt sogar folgende Aussage:

**Satz 5.4.18.** *Eine Norm $\|\cdot\|$ auf einem $\mathbb{K}$-Vektorraum $V$ wird genau dann von einen Skalarprodukt erzeugt, wenn die Parallelogrammgleichung gilt. In diesem Fall ist das Skalarprodukt gegeben durch*

$$\langle v, w\rangle = \begin{cases} \frac{1}{4}(\|v + w\|^2 - \|v - w\|^2 + i\|v - iw\|^2 - i\|v + iw\|^2) & \text{falls } \mathbb{K} = \mathbb{C} \\ \frac{1}{4}(\|v + w\|^2 - \|v - w\|^2) & \text{falls } \mathbb{K} = \mathbb{R} \end{cases} \tag{5.4.11}$$

**Lemma 5.4.19.** *Es sei $V$ ein $\mathbb{K}-$ Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Sind $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ konvergente Folgen in $V$ mit $x = \lim\limits_{n\to\infty} x_n$ und $y = \lim\limits_{n\to\infty} y_n$, dann ist $(\langle x_n, y_n\rangle)_{n\in\mathbb{N}}$ eine konvergente Folge in $\mathbb{K}$ mit*

$$\langle x, y\rangle = \lim_{n\to\infty} \langle x_n, y_n\rangle$$

*Beweis.* bleibt als Aufgabe. □

**Beispiel* 5.4.20.** Ist $d \in \mathbb{N}$, $p \in [1, \infty[$, so wird auf $\mathbb{K}^d$ durch

$$\left\| \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \right\|_p := [|x_1|^p + ... + |x_d|^p]^{\frac{1}{p}} \tag{5.4.12}$$

bzw. durch

$$\left\| \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \right\|_\infty := \max\{|x_1|, ..., |x_d|\} \tag{5.4.13}$$

eine Norm definiert. Für $p = 2$ ist $\|\cdot\|_2$ die vom Standardskalarprodukt induzierte Norm. In allen anderen Fällen fehlt zumindest noch der Beweis – bzw. im Fall $p \in \mathbb{R}\setminus\mathbb{Q}$ müssen wir auch noch $|x|^p$ definieren.

## 5.5 Limessuperior und Limesinferior reeller Folgen

**Lemma 5.5.1.** *Es sei $(K, \leq)$ ein geordneter Körper und $(a_n)_{n\in\mathbb{N}}$ eine Folge in $K$. Dann gibt es eine monotone Teilfolge $(a_{n_j})_{j\in\mathbb{N}}$ von $(a_n)_{n\in\mathbb{N}}$*

*Beweis.* Gibt es ein $k \in \mathbb{N}$ mit $a_k \geq a_n$ für alle $n \geq k$, so nennen wir $k$ eine Spitze für die Folge $(a_n)_{n\in\mathbb{N}}$.

- Gibt es für $(a_n)_{n\in\mathbb{N}}$ unendlich viele Spitzen, so bildet die Folge $(a_k)_{k\in\mathbb{N}}$ der Spitzen eine monoton fallende Teilfolge.
- Gibt es nur endlich viele Spitzen, so können wir $N \in \mathbb{N}$ wählen, so daß für $m \geq N$ das Element $m$ keine Spitze zu $(a_n)_{n\in\mathbb{N}}$ ist. Da $N$ keine Spitze ist, gibt es ein $N_1 > N$ mit $a_{N_1} > a_N$. Weil auch $N_1$ keine Spitze ist, gibt es $N_2 > N_1$ mit $a_{N_2} > a_{N_1}$,... und so findet man rekursiv eine monoton wachsende Teilfolge $(a_{N_j})_{j\in\mathbb{N}}$. □

**Satz 5.5.2** (Bolzano-Weierstraß). *Jede beschränkte Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ besitzt eine konvergente Teilfolge $(x_{n_j})_{j\in\mathbb{N}}$.*

*Beweis.* Nach Lemma 5.5.1 gibt es eine monotone Teilfolge $(x_{n_j})_{j\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$. Wegen $\{x_{n_j} : j \in \mathbb{N}\} \subseteq \{x_n : n \in \mathbb{N}\}$ ist auch die Teilfolge $(x_{n_j})_{j\in\mathbb{N}}$ in $\mathbb{R}$ beschränkt, und konvergiert daher in $\mathbb{R}$ nach Lemma 5.2.4 bzw. Korollar 5.2.5. □

**Satz 5.5.3.** *Es sei $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$, dann existiert*

$$\limsup_{n\to\infty} x_n := \lim_{n\to\infty} \left(\sup\{x_k : k \geq n\}\right) \tag{5.5.1}$$

$$\liminf_{n\to\infty} x_n := \lim_{n\to\infty} \left(\inf\{x_k : k \geq n\}\right) \tag{5.5.2}$$

*und für $H := \big\{x \in \mathbb{R} : x$ ist Häufungswert von $(x_n)_{n\in\mathbb{N}}\big\}$ gilt:*

$$\limsup_{n\to\infty} x_n = \inf\Big\{\sup\{x_k : k \geq n\} : n \in \mathbb{N}\Big\} = \max(H) \tag{5.5.3}$$

$$\liminf_{n\to\infty} x_n = \sup\Big\{\inf\{x_k : k \geq n\} : n \in \mathbb{N}\Big\} = \min(H) \tag{5.5.4}$$

*Beweis.* Zur Abkürzung schreibe $X_n := \{x_k : k \geq n\}$, $y_n := \sup(X_n)$, $Y := \{y_n : n \in \mathbb{N}\}$. Wegen $X_m \subseteq X_n$ für $m \geq n$ ist die Folge $(y_n)_{n\in\mathbb{N}}$ monoton fallend und wegen $\inf(X_1) \leq x_n \leq y_n$ nach unten beschränkt. Nach Korollar 5.2.5 existiert dann $\lim\limits_{n\to\infty} y_n$ und wir bekommen $\limsup\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} y_n = \inf(Y)$, was die Existenz von $\limsup\limits_{n\in\mathbb{N}} x_n$ und die erste Gleichheit in (5.5.3) zeigt. Wir definieren nun rekursiv eine Teilfolge $(x_{m_j})_{j\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ mit $\gamma := \inf(Y) = \lim\limits_{j\to\infty} x_{m_j}$. Nach Korollar 4.1.7 angewandt auf $\gamma = \inf(Y)$ ist $\{k \in \mathbb{N} : y_k \in [\gamma, \gamma + 1[\} \neq \emptyset$ und wir wählen daher $n_1 := \min\{k \in \mathbb{N} : y_k \in [\gamma, \gamma + 1[\}$. Dann ist $\{l \in \mathbb{N} : x_l \in ]y_{n_1} - 1, y_{n_1}] \cap X_{n_1}\} \neq \emptyset$, denn $y_{n_1} = \sup(X_{n_1})$; wir wählen nun $m_1 := \min\{l \in \mathbb{N} : x_l \in ]y_{n_1} - 1, y_{n_1}] \cap X_{n_1}\}$. Sind nun $n_1, ..., n_j$ und $m_1, ...., m_j$ konstruiert, dann ist – da $(y_n)_{n\in\mathbb{N}}$ monoton fallend ist – $\{k \in \mathbb{N} : k > m_j \geq n_j, y_k \in [\gamma, \gamma + \frac{1}{j+1}[\} \neq \emptyset$. Wähle also

$$\begin{aligned} n_{j+1} &:= \min\{k \in \mathbb{N} : k > n_j, y_k \in [\gamma, \gamma + \frac{1}{j+1}[\} \\ m_{j+1} &:= \min\{l \in \mathbb{N} : x_l \in ]y_{n_{j+1}} - \frac{1}{j+1}, y_{n_{j+1}}] \cap X_{n_{j+1}}\}, \end{aligned}$$

dann ist $(n_j)_{j\in\mathbb{N}}$ eine streng monoton steigende Folge, $m_j \geq n_j$ und $\gamma - \frac{1}{j} < x_{m_j} < \gamma + \frac{1}{j}$. Somit ist $\gamma = \lim_{j\to\infty} x_{m_j}$ ein Häufungswert von $(x_n)_{n\in\mathbb{N}}$. Zu $z \in H$ gibt es eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ mit $z = \lim_{k\to\infty} x_{n_k}$. Da $(y_n)_{n\in\mathbb{N}}$ konvergiert, konvergiert auch die Teilfolge $(y_{n_k})_{k\in\mathbb{N}}$ und aus $x_{n_k} \leq y_{n_k}$ folgt dann $z = \lim_{k\to\infty} x_{n_k} \leq \lim_{k\to\infty} y_{n_k} = \lim_{n\to\infty} y_n = \gamma$, dh. $\gamma$ ist obere Schranke von $H$ und damit $\gamma = \max(H)$. Dies zeigt die 2. Gleichheit in (5.5.3). Die Existenz von $\liminf_{n\to\infty} x_n$ und (5.5.4) zeigt man analog. □

**Lemma 5.5.4.** *Es sei $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$, dann ist $\gamma = \limsup_{n\to\infty} x_n$ durch die beiden Bedingungen*

*a) $\gamma$ ist Häufungswert von $(x_n)_{n\in\mathbb{N}}$.*

*b) Für alle $y > \gamma$ gibt es ein $N \in \mathbb{N}$, so daß für $n \geq N$ gilt: $x_n < y$.*

*eindeutig festgelegt.*

*Beweis.* Nach Satz 5.5.3 existiert $\gamma = \limsup_{n\to\infty} x_n$ und ist der größte Häufungswert von $(x_n)_{n\in\mathbb{N}}$, daher gilt a) für $\gamma = \limsup_{n\to\infty} x_n$. Ist b) nicht erfüllt, so gibt es ein $y > \gamma$, so daß für alle $N \in \mathbb{N}$ ein $n \geq N$ existiert mit $x_n \geq y$. Wir finden daher eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ mit $x_{n_k} \geq y$ für alle $k \in \mathbb{N}$. Nach dem Satz von Bolzano-Weierstraß gibt es eine konvergente Teilfolge $(y_{n_j})_{j\in\mathbb{N}}$ von $(x_{n_k})_{k\in\mathbb{N}}$ und nach Lemma 5.2.2 gilt $\beta := \lim_{j\to\infty} y_{n_j} \geq y$ und somit ist $\beta > \gamma$ Häufungswert von $(x_n)_{n\in\mathbb{N}}$ im Widerspruch zu $\gamma$ größter Häufungswert. Wir haben also damit gezeigt, daß $\gamma = \limsup_{n\to\infty} x_n$ die Bedingungen a) und b) erfüllt. Angenommen es gibt $\beta, \gamma \in \mathbb{R}$ mit $\beta < \gamma$, die die Eigenschaften a) und b) erfüllen. Wähle $x \in \mathbb{R}$ mit $\beta < x < \gamma$, dann gibt es – da $\beta$ Bedingung b) erfüllt – ein $N \in \mathbb{N}$ mit $x_n < x$ für alle $n \geq N$. Aber dann kann $\gamma$ kein Häufungswert von $(x_n)_{n\in\mathbb{N}}$ sein im Widerspruch zur Voraussetzung. □

**Korollar 5.5.5.** *Es sei $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$, dann ist $\gamma = \liminf_{n\to\infty} x_n$ durch die beiden Bedingungen*

*a) $\gamma$ ist Häufungswert von $(x_n)_{n\in\mathbb{N}}$.*

*b) Für alle $y < \gamma$ gibt es ein $N \in \mathbb{N}$, so daß für $n \geq N$ gilt: $x_n > y$.*

*eindeutig festgelegt.*

**Lemma 5.5.6.** *Eine Folge* $(x_n)_{n\in\mathbb{N}}$ *in* $\mathbb{R}$ *konvergiert genau dann wenn* $\liminf\limits_{n\to\infty} x_n$ *und* $\limsup\limits_{n\to\infty} x_n$ *existieren und* $\liminf\limits_{n\to\infty} x_n = \limsup\limits_{n\to\infty} x_n$ *gilt. In diesem Fall ist*

$$\lim_{n\to\infty} x_n = \liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n.$$

*Beweis.* Ist $(x_n)_{n\in\mathbb{N}}$ nicht beschränkt, so kann $(x_n)_{n\in\mathbb{N}}$ nach Lemma 5.1.8 nicht konvergieren. Es sei also $(x_n)_{n\in\mathbb{N}}$ beschränkt, dann existieren $\liminf\limits_{n\to\infty} x_n$ und $\limsup\limits_{n\to\infty} x_n$ nach Satz 5.5.3. Nach dem Satz von Bolzano-Weierstraß ist die Menge

$$H := \{x \in \mathbb{R} : x \text{ ist Häufungswert von } (x_n)_{n\in\mathbb{N}}\} \neq \emptyset,$$

daher gilt

$$\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n$$

nach (5.5.3) und (5.5.4).

„$\Rightarrow$" Es sei $(x_n)_{n\in\mathbb{N}}$ konvergent und $x = \lim\limits_{n\to\infty} x_n$, dann ist $x$ nach Lemma 5.1.7 der einzige Häufungswert von $(x_n)_{n\in\mathbb{N}}$, also

$$\liminf_{n\to\infty} x_n = \min(H) = x = \max(H) = \limsup_{n\to\infty} x_n$$

„$\Leftarrow$" Es sei $\liminf\limits_{n\to\infty} x_n = \limsup\limits_{n\to\infty} x_n =: x$, dann ist $\min(H) = \max(H) = x$. Ist $\varepsilon > 0$, dann ist $x + \varepsilon > x$, daher gibt es nach Lemma 5.5.4 ein $N(\varepsilon) \in \mathbb{N}$, so daß für $n \geq N(\varepsilon)$ gilt: $x_n < x + \varepsilon$. Nach Korollar 5.5.5 gibt es $M(\varepsilon) \in \mathbb{N}$, so daß für $n \geq M(\varepsilon)$ gilt: $x - \varepsilon < x_n$, damit gilt: $x - \varepsilon < x_n < x + \varepsilon$ für alle $n \geq \max\{N(\varepsilon), M(\varepsilon)\}$, dh. $x = \lim\limits_{n\to\infty} x_n$.

□

## 5.6 Die erweiterte reelle Zahlengerade $\widehat{\mathbb{R}}$

**Lemma 5.6.1.** $\begin{array}{rcl} f : \mathbb{R} & \to & ]-1,1[ \\ x & \mapsto & \frac{x}{1+|x|} \end{array}$ *ist streng monoton steigend und bijektiv.*

*Beweis.* Wegen $|x| < 1 + |x|$ ist $\left|\frac{x}{1+|x|}\right| < 1$, also $]-1,1[$ als Wertevorrat möglich. Sind $x, y \in \mathbb{R}$ mit $x < y$, dann:

- für $0 \leq x < y$ folgt $x + xy = x(1+y) < y(1+x) = y + xy$ und dann wegen $1+x, 1+y \geq 1 > 0$ auch $f(x) = \frac{x}{1+x} < f(y) = \frac{y}{1+y}$.

- für $x < 0 \leq y$ folgt $x < y - 2xy$, also $x + xy = x(1+y) < y(1-x) = y - xy$ und dann wegen $1-x, 1+y \geq 1 > 0$ auch $f(x) = \frac{x}{1-x} < \frac{y}{1+y} = f(y)$.

- für $x < y \leq 0$ folgt $x - xy = x(1-y) < y(1-x) = y - xy$ und wegen $1-x, 1-y \geq 1 > 0$ folgt $f(x) = \frac{x}{1-x} < \frac{y}{1-y} = f(y)$, was zeigt, daß $f$ streng monoton steigend ist. Ist $y \in ]-1,1[$, dann ist $1-|y|>0$, also $\begin{array}{rcl} g: ]-1,1[ & \to & \mathbb{R} \\ y & \mapsto & \frac{y}{1-|y|} \end{array}$ wohldefiniert und es gilt:

$$\begin{array}{rcl} f\circ g: ]-1,1[ & \to & ]-1,1[ \\ y & \mapsto & f(\frac{y}{1-|y|}) = \frac{\frac{y}{1-|y|}}{1+|\frac{y}{1-|y|}|} = y \end{array}$$

und

$$\begin{array}{rcl} g\circ f: \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & g(\frac{x}{1+|x|}) = \frac{\frac{x}{1+|x|}}{1-|\frac{x}{1+|x|}|} = x \end{array}$$

also ist $f\circ g = \mathrm{id}_{]-1,1[}$ und $g\circ f = \mathrm{id}_{\mathbb{R}}$ und daher ist $f$ bijektiv. $\square$

**Lemma 5.6.2.** *Auf der erweiterten reellen Zahlengeraden* $\widehat{\mathbb{R}} := \mathbb{R}\cup\{-\infty,\infty\}$ *wird durch* $\begin{array}{rcl} F: \widehat{\mathbb{R}} & \to & [-1,1] \\ x & \mapsto & \begin{cases} \frac{x}{1+|x|} & \text{für } x\in\mathbb{R} \\ -1 & \text{für } x=-\infty \\ 1 & \text{für } x=\infty \end{cases} \end{array}$ *eine bijektive Abbildung definiert.*

- *Für* $x,y\in\widehat{\mathbb{R}}$ *wird durch*

$$x \le y \text{ genau dann wenn } F(x)\le F(y) \tag{5.6.1}$$

  *eine totale Ordnungsrelation auf* $\widehat{\mathbb{R}}$ *definiert. Diese Ordnungsrelation setzt die Ordnung von* $\mathbb{R}$ *auf* $\widehat{\mathbb{R}}$ *fort und erfüllt* $-\infty < x < \infty$ *für alle* $x\in\mathbb{R}$.

- $\begin{array}{rcl} d: \widehat{\mathbb{R}}\times\widehat{\mathbb{R}} & \to & [0,\infty[ \\ (x,y) & \mapsto & |F(x)-F(y)| \end{array}$ *ist eine Metrik auf* $\widehat{\mathbb{R}}$ *und macht damit* $(\widehat{\mathbb{R}},d)$ *zu einem metrischen Raum.*

*Beweis.* Da die Abbildung $\begin{array}{rcl} f: \mathbb{R} & \to & ]-1,1[ \\ x & \mapsto & \frac{x}{1+|x|} \end{array}$ nach Lemma 5.6.1 bijektiv ist, gilt dies auch für $F$. Da außerdem $[-1,1]$ total geordnet ist, definiert (5.6.1) eine totale Ordnung auf $\widehat{\mathbb{R}}$ und da $F|_{\mathbb{R}} = f$ streng monoton steigend ist, setzt diese Definition die gegebene Ordnung von $\mathbb{R}$ auf $\widehat{\mathbb{R}}$ fort. $d$ definiert eine Metrik auf $\widehat{\mathbb{R}}$; denn sind $x,y,z\in\widehat{\mathbb{R}}$ so gilt:

- Wegen positiver Definitheit des Absolutbetrags auf $\mathbb{R}$ ist $d(x,y)=|F(x)-F(y)|=0$ genau dann wenn $F(x)=F(y)$ ist, was wegen Bijektivität von $F$ zu $x=y$ äquivalent ist.

- $d(x,y)=|F(x)-F(y)|=d(y,x)$ folgt aus der Symmetrie von $|\cdot|$.

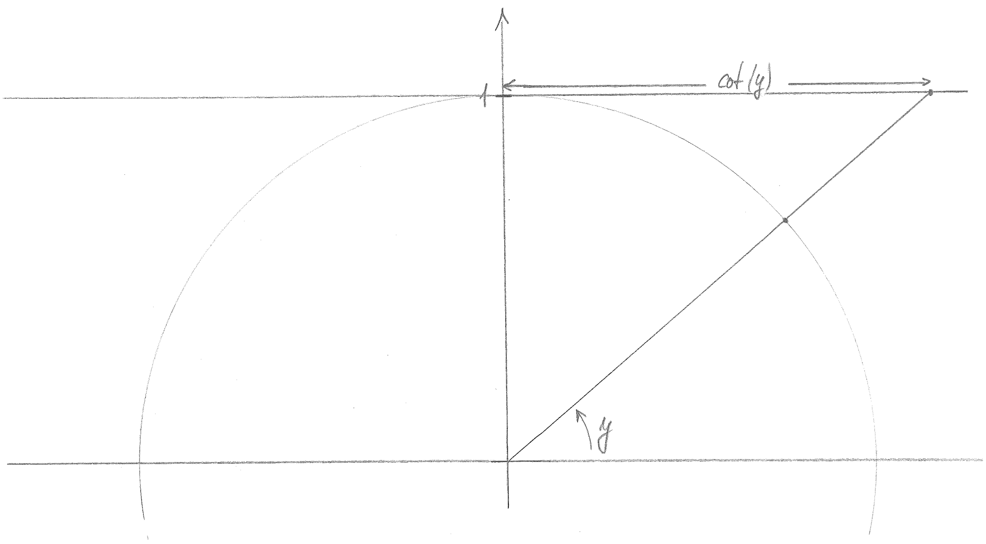
- Aus der Dreiecksungleichung von $|\cdot|$ folgt:

$$\begin{array}{rcl} d(x,z) & = & |F(x)-F(z)| = |F(x)-F(y)+F(y)-F(z)| \\ & \le & |F(x)-F(y)|+|F(y)-F(z)| = d(x,y)+d(y,z). \end{array}$$

$\square$

**Bemerkung 5.6.3.**

- Die Umkehrfunktion $\begin{array}{rcl} g : ]-1,1[ & \to & \mathbb{R} \\ y & \mapsto & \frac{y}{1-|y|} \end{array}$ von $f$ macht etwas Ähnliches wie $\begin{array}{rcl} \cot : ]0,\pi[ & \to & \mathbb{R} \\ y & \mapsto & \cot(y) \end{array}$. Bei cot kann man es sich geometrisch leicht klar machen, was passiert;




mit $g$ und $f$ läßt sich leichter rechnen.

- Eine häufig benutzte Schreibweise für reelle Folgen $(x_n)_{n\in\mathbb{N}}$ wie $\lim\limits_{n\to\infty} x_n = \infty$ ist in $(\widehat{\mathbb{R}}, d)$ jetzt klar definiert: Faßt man wegen $\mathbb{R} \subseteq \widehat{\mathbb{R}}$ die reelle Folge $(x_n)_{n\in\mathbb{N}}$ als eine Folge in $\widehat{\mathbb{R}}$ auf, so gilt $\infty = \lim\limits_{n\to\infty} x_n$ im metrischen Raum $(\widehat{\mathbb{R}}, d)$ genau dann, wenn
$$0 = \lim_{n\to\infty} |F(x_n) - 1|$$
ist. Da $F$ streng monoton steigend ist, läßt sich das auch ohne „Umweg“ über $\widehat{\mathbb{R}}$ ganz allein in $\mathbb{R}$ formulieren: Für jedes $r > 0$ gibt es $N = N(r) \in \mathbb{N}$ mit $x_n \geq r$ für alle $n \geq N(r)$.

- Analog bedeutet $x_n \overset{n\to\infty}{\longrightarrow} -\infty$, daß
$$0 = \lim_{n\to\infty} |F(x_n) + 1|$$
ist, oder für jedes $s < 0$ gibt es $N = N(s) \in \mathbb{N}$ mit $x_n \leq s$ für alle $n \geq N(s)$.

**Lemma 5.6.4.** *Sind $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ Folgen in $\widehat{\mathbb{R}}$ mit $x := \lim\limits_{n\to\infty} x_n$, $y := \lim\limits_{n\to\infty} y_n$ und $x_n \leq y_n$ für alle $n \in \mathbb{N}$, dann ist auch $x \leq y$*

*Beweis.* Nach Definition bedeutet $x_n \leq y_n$ gerade $F(x_n) \leq F(y_n)$ und genauso ist $x = \lim\limits_{n\to\infty} x_n$ äquivalent zu $\lim\limits_{n\to\infty} d(x, x_n) = 0$ bzw. $F(x) = \lim\limits_{n\to\infty} F(x_n)$ in $\mathbb{R}$. Für die beiden konvergenten reellen Folgen $(F(x_n))_{n\in\mathbb{N}}$ bzw. $(F(y_n))_{n\in\mathbb{N}}$ folgt dann aber $F(x) \leq F(y)$ nach Lemma 5.2.2, also $x \leq y$. □

**Lemma 5.6.5.** *Es sei $x \in \mathbb{R}$. Eine Folge $(x_n)_{n\in\mathbb{N}}$ reeller Zahlen konvergiert genau dann in $\mathbb{R}$ gegen $x$, wenn die Folge in $(\widehat{\mathbb{R}}, d)$ gegen $x$ konvergiert.*

*Beweis.*

„⇒" Gilt $x = \lim\limits_{n\to\infty} x_n$ in $\mathbb{R}$, dann folgt wegen $1 + |x_n| \geq 1$ aus den Rechenregeln für konvergente Folgen reeller Zahlen in Lemma 5.2.1:

$$F(x_n) = f(x_n) = \frac{x_n}{1 + |x_n|} \overset{n\to\infty}{\longrightarrow} \frac{x}{1 + |x|} = f(x) = F(x),$$

dh. $\lim\limits_{n\to\infty} d(x, x_n) = \lim\limits_{n\to\infty} |F(x_n) - F(x)| = 0$ oder $x = \lim\limits_{n\to\infty} x_n$ in $(\widehat{\mathbb{R}}, d)$.

„⇐" Ist $x \in \mathbb{R}$ und $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ mit

$$\lim_{n\to\infty} d(x, x_n) = \lim_{n\to\infty} |F(x) - F(x_n)| = 0. \tag{5.6.2}$$

Wegen $x \in \mathbb{R}$ ist $F(x) = f(x) \in ]-1, 1[$, also $\varepsilon := \frac{1}{2}\min\{|f(x) + 1|, |f(x) - 1|\} > 0$. Wegen $x_n \in \mathbb{R}$ und (5.6.2) gibt es ein $N \in \mathbb{N}$ mit $f(x_n) \in ]f(x) - \varepsilon, f(x) + \varepsilon[$ für alle $n \geq N$. Da $g = f^{-1}$ streng monoton steigend ist, folgt

$$x_n = (g \circ f)(x_n) \in ]g(f(x) - \varepsilon), g(f(x) + \varepsilon)[$$

für alle $n \geq N$, daher ist die Folge $(x_n)_{n\in\mathbb{N}}$ beschränkt in $\mathbb{R}$. Da

$$\begin{aligned} |F(x) - F(x_n)| &= |f(x) - f(x_n)| = \frac{|x(1 + |x_n|) - x_n(1 + |x|)|}{(1 + |x|)(1 + |x_n|)} \\ &= \left| \frac{x - x_n}{(1 + |x|)(1 + |x_n|)} + f(x)|f(x_n)| - f(x_n)|f(x)| \right| \overset{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

und $f(x) = \lim\limits_{n\to\infty} f(x_n)$ ist, ergibt sich $\left|\frac{x - x_n}{(1+|x|)(1+|x_n|)}\right| \overset{n\to\infty}{\longrightarrow} 0$ und da $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ beschränkt ist folgt daraus $\lim\limits_{n\to\infty} |x - x_n| = 0$.

□

**Definition 5.6.6.** *Es sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\widehat{\mathbb{R}}$, dann heißt*

$$\overline{\lim_{n\to\infty}}\, x_n := F^{-1}\left(\limsup_{n\to\infty} F(x_n)\right) \tag{5.6.3}$$

*der* **Limessuperior** *und*

$$\underline{\lim_{n\to\infty}}\, x_n := F^{-1}\left(\liminf_{n\to\infty} F(x_n)\right) \tag{5.6.4}$$

*der* **Limesinferior** *der Folge $(x_n)_{n\in\mathbb{N}}$.*

**Bemerkung 5.6.7.**

- Weil für jede Folge $(x_n)_{n\in\mathbb{N}}$ in $\widehat{\mathbb{R}}$ die Folge $(F(x_n))_{n\in\mathbb{N}}$ eine Folge in $[-1,1]$ ist, existieren $\limsup\limits_{n\to\infty} F(x_n) \in [-1,1]$ und $\liminf\limits_{n\to\infty} F(x_n) \in [-1,1]$ nach Satz 5.5.3. Daher sind in (5.6.3) bzw. (5.6.4) die Ausdrücke $F^{-1}\left(\limsup\limits_{n\to\infty} F(x_n)\right)$ und $F^{-1}\left(\liminf\limits_{n\to\infty} F(x_n)\right)$ wohldefiniert.

- Ist $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$, so gilt
$$\limsup_{n\to\infty} x_n = \overline{\lim_{n\to\infty}}\, x_n \qquad \text{und} \qquad \liminf_{n\to\infty} x_n = \underline{\lim_{n\to\infty}}\, x_n \tag{5.6.5}$$
Denn ist $z \in \mathbb{R}$ ein Häufungswert von $(x_n)_{n\in\mathbb{N}}$ und $(x_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge von $(x_n)_{n\in\mathbb{N}}$ mit $z = \lim\limits_{k\to\infty} x_{n_k}$, so folgt wegen $1 + |x_{n_k}| \geq 1$ aus den Rechenregeln für konvergente reelle Folgen: $f(x_{n_k}) = \frac{x_{n_k}}{1+|x_{n_k}|} \xrightarrow{k\to\infty} \frac{z}{1+|z|} = f(z)$, dh. $f(z)$ ist Häufungswert von $(f(x_n))_{n\in\mathbb{N}}$. Da $f$ streng monoton steigend und bijektiv ist, gilt für $x := \limsup\limits_{n\to\infty} x_n$ – was gemäß Satz 5.5.3 größter Häufungswert von $(x_n)_{n\in\mathbb{N}}$ ist –
$$f(x) = f(\limsup_{n\to\infty} x_n) = \limsup_{n\to\infty} f(x_n).$$
Daraus folgt
$$\limsup_{n\to\infty} x_n = x = f^{-1}(f(x)) = F^{-1}(\limsup_{n\to\infty} F(x_n)) = \overline{\lim_{n\to\infty}}\, x_n.$$

- Ist $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$, die nicht nach oben beschränkt ist, dann ist $\overline{\lim\limits_{n\to\infty}}\, x_n = \infty$. Wähle dazu eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ mit $x_{n_k} \geq k$, dann ist
$$\limsup_{n\to\infty} F(x_n) \geq \lim_{k\to\infty} F(x_{n_k}) \geq \lim_{k\to\infty} F(k) = \lim_{k\to\infty} \frac{k}{1+k} = 1$$
also $\limsup\limits_{n\to\infty} F(x_n) = 1$ und somit $\overline{\lim\limits_{n\to\infty}}\, x_n = F^{-1}(1) = \infty$.

- Ist $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\widehat{\mathbb{R}}$ mit $\underline{\lim\limits_{n\to\infty}}\, x_n = \overline{\lim\limits_{n\to\infty}}\, x_n =: x \in \widehat{\mathbb{R}}$, dann gilt $x = \lim\limits_{n\to\infty} x_n$ in $(\widehat{\mathbb{R}}, d)$, denn aus
$$x = F^{-1}\left(\liminf_{n\to\infty} F(x_n)\right) = F^{-1}\left(\limsup_{n\to\infty} F(x_n)\right)$$
folgt $\liminf\limits_{n\to\infty} F(x_n) = \limsup\limits_{n\to\infty} F(x_n)$, woraus die Existenz von
$$y := F(x) = \lim_{n\to\infty} F(x_n) = \liminf_{n\to\infty} F(x_n) = \limsup_{n\to\infty} F(x_n) \in [-1,1]$$
nach Lemma 5.5.6 folgt, dh. $\lim\limits_{n\to\infty} |F(x_n) - y| = 0$ oder
$$\lim_{n\to\infty} d(x_n, F^{-1}(y)) = \lim_{n\to\infty} d(x_n, x) = 0.$$

**Bemerkung 5.6.8. Vorsicht:** Auch wenn sich, wie eben bemerkt einige Rechenregeln für konvergente Folgen von $\mathbb{R}$ auf $\widehat{\mathbb{R}}$ übertragen, so gilt das nicht für alle. Auf $\widehat{\mathbb{R}}$ läßt sich keine Addition definieren, so daß sich diese Definition mit den Regeln für die Summen von Grenzwerten reeller Folgen verträgt. Etwa $(+\infty) + (-\infty)$ kann man so nicht definieren: Ein Beispiel:

- Die Folge $(a_n)_{n\in\mathbb{N}} = (n)_{n\in\mathbb{N}}$ hat in $\widehat{\mathbb{R}}$ den Grenzwert $+\infty$ und $(b_n)_{n\in\mathbb{N}} := (-n)_{n\in\mathbb{N}}$ hat in $\widehat{\mathbb{R}}$ den Grenzwert $-\infty$, denn

$$d(n,\infty) = \left|\frac{n}{1+n} - 1\right| = \frac{1}{1+n} \overset{n\to\infty}{\longrightarrow} 0$$

$$d(-n,-\infty) = \left|\frac{-n}{1+n} + 1\right| = \frac{1}{1+n} \overset{n\to\infty}{\longrightarrow} 0$$

  $(a_n + b_n)_{n\in\mathbb{N}} := (n-n)_{n\in\mathbb{N}}$ hat als konstante Folge den Grenzwert 0.

- Die Folgen $(c_n)_{n\in\mathbb{N}} = (\sqrt{n+\sqrt{n}})_{n\in\mathbb{N}}$ und $(d_n)_{n\in\mathbb{N}} = (\sqrt{n})_{n\in\mathbb{N}}$ haben den Grenzwert $\infty$ in $\widehat{\mathbb{R}}$, denn für jedes $\varepsilon > 0$ ist

$$d(\sqrt{n},\infty) = \left|\frac{\sqrt{n}}{1+\sqrt{n}} - 1\right| = \frac{1}{1+\sqrt{n}} < \varepsilon$$

$$d(\sqrt{n+\sqrt{n}},\infty) = \left|\frac{\sqrt{n+\sqrt{n}}}{1+\sqrt{n+\sqrt{n}}} - 1\right| = \frac{1}{1+\sqrt{n+\sqrt{n}}} \le \frac{1}{1+\sqrt{n}} < \varepsilon$$

  für alle $n > \frac{1}{\varepsilon^2}$. Weil $1 + \frac{1}{\sqrt{n}} \ge 1$ ist, gilt $1 \le \sqrt{1+\frac{1}{\sqrt{n}}} \le 1 + \frac{1}{\sqrt{n}}$ für jedes $n \in \mathbb{N}$.

  Da wie eben gesehen, $0 = \lim\limits_{n\to\infty} \frac{1}{\sqrt{n}}$ ist, folgt $1 = \lim\limits_{n\to\infty} \sqrt{1+\frac{1}{\sqrt{n}}}$ nach Lemma 5.2.3.

  Damit ergibt sich für

$$\begin{aligned} c_n - d_n &= \sqrt{n+\sqrt{n}} - \sqrt{n} = \frac{(\sqrt{n+\sqrt{n}} - \sqrt{n})(\sqrt{n+\sqrt{n}} + \sqrt{n})}{\sqrt{n+\sqrt{n}} + \sqrt{n}} \\ &= \frac{n + \sqrt{n} - n}{\sqrt{n+\sqrt{n}} + \sqrt{n}} = \frac{1}{1+\sqrt{1+\frac{1}{\sqrt{n}}}} \overset{n\to\infty}{\longrightarrow} \frac{1}{2} \end{aligned}$$

# Kapitel 6

# Reihen

## 6.1 Reihen

**Definition 6.1.1.** *Es sei* $(V, \|\cdot\|)$ *ein normierter* $\mathbb{K}$*-Vektorraum,* $\begin{array}{rcl} d: V \times V & \to & [0,\infty[ \\ (v,w) & \mapsto & \|v-w\| \end{array}$ *die von* $\|\cdot\|$ *induzierte Metrik und* $(x_k)_{k\in\mathbb{N}}$ *eine Folge in* $V$*. Eine Folge der Form* $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ *heißt eine* **Reihe***. Die* $x_k$ *heißen die* **Reihenglieder** *und*

$$s_n := \sum_{k=1}^{n} x_k$$

*heißen die* **Partialsummen** *der Reihe* $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$*. Die Reihe* $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ *heißt* **konvergent***, wenn* $\lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=1}^{n} x_k$ *in* $(V,d)$ *existiert. In diesem Fall schreibe*

$$\sum_{k=1}^{\infty} x_k := \lim_{n\to\infty} \sum_{k=1}^{n} x_k. \tag{6.1.1}$$

**Bemerkung 6.1.2.**

- Ist $z \in \mathbb{Z}$ und $(x_k)_{k\geq z}$ gegeben, kann man genauso die Partialsummenfolge $(s_n)_{n\geq z}$ mit $s_n := \sum_{k=z}^{n} x_k$ und die Reihe $\left(\sum_{k=z}^{n} x_k\right)_{n\geq z}$ definieren. Die folgenden Ergebnisse sind für $z = 1$ formuliert, gelten aber auch – nach Addition oder Subtraktion von endlich vielen Gliedern der Reihe – für jedes andere gegebene $z \in \mathbb{Z}$ – für das man $(x_n)_{n\geq z}$ und die zugehörige Reihe definieren kann.

- In der Literatur sind die Bezeichnungen nicht ganz einheitlich. Oft wird $\sum_{k=1}^{\infty} x_k$ als Schreibweise sowohl für die Folge $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ der Partialsummen, als auch für deren Grenzwert $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} x_k$ verwendet.

**Beispiel 6.1.3.**

a) Ist $z \in \mathbb{C}$, $|z| < 1$, so existiert der Grenzwert $\sum\limits_{k=0}^{\infty} z^k$ der **geometrischen Reihe** $\left(\sum\limits_{k=1}^{n} z^k\right)_{n\in\mathbb{N}}$ und es gilt

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1-z} \tag{6.1.2}$$

*Beweis.* Für jede Partialsumme können wir für $z \neq 1$ wie in Beispiel 2.1.4

$$\sum_{k=0}^{n} z^k = 1 + z + ... + z^n = \frac{1-z^{n+1}}{1-z}$$

zeigen. Wegen $|z| < 1$ gilt $|z^{n+1}| = |z|^{n+1} \overset{n\to\infty}{\longrightarrow} 0$, daher existiert

$$\sum_{k=0}^{\infty} z^k = \lim_{n\to\infty} \sum_{k=0}^{n} z^k = \lim_{n\to\infty} \frac{1-z^{n+1}}{1-z} = \frac{1}{1-z}.$$ □

b) Die harmonische Reihe $\left(\sum\limits_{k=1}^{n} \frac{1}{k}\right)_{n\in\mathbb{N}}$ konvergiert nicht.

*Beweis.* Betrachte die Teilfolge $(s_{2^n})_{n\in\mathbb{N}}$ der Folge der Partialsummen. Dann ist

$$s_{2^n} = \sum_{k=1}^{2^n} \frac{1}{k} = 1 + \frac{1}{2} + \underbrace{(\frac{1}{3} + \frac{1}{4})}_{\geq \frac{1}{2}} + \underbrace{(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8})}_{\geq \frac{1}{2}} + ... + \underbrace{(\frac{1}{2^{n-1}+1} + ... + \frac{1}{2^n})}_{\geq \frac{1}{2}} \geq 1 + \frac{n}{2}$$

und damit die Partialsummenfolge $(\sum\limits_{k=1}^{n} \frac{1}{k})_{n\in\mathbb{N}}$ nicht beschränkt, also auch nicht konvergent. □

c) $b-$adische Entwicklung einer reellen Zahl in Beispiel 5.2.7

**Bemerkung 6.1.4.** Es sei $(V, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum.

- Ist $(x_k)_{k\in\mathbb{N}}$ eine Folge in $V$, die nicht gegen $\mathbf{0}$ konvergiert, so konvergiert die Reihe $\left(\sum\limits_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ nicht.

*Beweis.* Ist $(x_k)_{k\in\mathbb{N}}$ eine Folge, die nicht gegen $\mathbf{0}$ konvergiert, so gilt (indem man die Definition einer Nullfolge verneint): Es gibt ein $\varepsilon > 0$ und für alle $N \in \mathbb{N}$ gibt es ein $n \geq N$ mit $\|x_n\| \geq \varepsilon$. Das heißt $(x_k)_{k\in\mathbb{N}}$ besitzt eine (unendliche) Teilfolge $(x_{k_l})_{l\in\mathbb{N}}$ mit $\|x_{k_l}\| \geq \varepsilon$. Damit ist die Partialsummenfolge $(s_n)_{n\in\mathbb{N}} = \left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ wegen $x_{k_l} = s_{k_l} - s_{k_l-1}$ und $\|x_{k_l}\| \geq \varepsilon$ keine Cauchyfolge, daher ist $(s_n)_{n\in\mathbb{N}}$ nach Lemma 5.1.6 auch nicht konvergent. □

- Wie die Anwendung der Definition einer Cauchyfolge auf die Folge der Partialsummen oder auch das Beispiel der harmonischen Reihe zeigt, **reicht es nicht** zu zeigen, daß $(x_k)_{k\in\mathbb{N}}$ eine Nullfolge ist, um die Konvergenz der Partialsummenfolge $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ zu beweisen. Nur wenn die Glieder $(x_k)_{k\in\mathbb{N}}$ einer Reihe **keine Nullfolge bilden**, hat die **Reihe** $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ **keine Chance zu konvergieren**.

**Lemma 6.1.5.** *Es sei $(V, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum, $\lambda \in \mathbb{K}$, $\left(\sum_{k=1}^{n} a_k\right)_{n\in\mathbb{N}}$ und $\left(\sum_{k=1}^{n} b_k\right)_{n\in\mathbb{N}}$ seien konvergente Reihen in $V$. Dann konvergiert in $V$ auch die Reihe $\left(\sum_{k=1}^{n}(a_k + \lambda b_k)\right)_{n\in\mathbb{N}}$ und es gilt:*

$$\sum_{k=1}^{\infty}(a_k + \lambda b_k) = \sum_{k=1}^{\infty} a_k + \lambda \sum_{k=1}^{\infty} b_k \tag{6.1.3}$$

*Beweis.* Da nach Voraussetzung die beiden Grenzwerte $x := \sum_{k=1}^{\infty} a_k$ und $y := \sum_{k=1}^{\infty} b_k$ existieren, gibt es für jedes $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ mit

$$\left\|x - \sum_{k=1}^{n} a_k\right\| < \frac{\varepsilon}{2(1+|\lambda|)} \quad \text{und} \quad \left\|y - \sum_{k=1}^{n} b_k\right\| < \frac{\varepsilon}{2(1+|\lambda|)} \quad \text{für alle } n \geq N(\varepsilon)$$

Nach den Regeln für Normen folgt:

$$\left\|x + \lambda y - \left(\sum_{k=1}^{n} a_k + \lambda \sum_{k=1}^{n} b_k\right)\right\| \leq \left\|x - \sum_{k=1}^{n} a_k\right\| + |\lambda| \left\|y - \sum_{k=1}^{n} b_k\right\| < \varepsilon$$

für alle $n \geq N(\varepsilon)$, dh. (6.1.3) ist gezeigt. □

**Lemma 6.1.6.** *Es sei $(V, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum und $(x_k)_{k\in\mathbb{N}}$ eine Folge in $V$, dann ist die* **Teleskopreihe** $\left(\sum_{k=1}^{n}(x_k - x_{k+1})\right)_{n\in\mathbb{N}}$ *genau dann konvergent, wenn*

$(x_k)_{k\in\mathbb{N}}$ *konvergiert. In diesem Fall gilt:*

$$\sum_{k=1}^{\infty}(x_k - x_{k+1}) = x_1 - \lim_{k\to\infty} x_k.$$

*Beweis.* Wegen $\sum_{k=1}^{n}(x_k - x_{k+1}) = x_1 - x_{n+1}$ braucht man nur in die Definitionen einzusetzen. □

**Beispiel 6.1.7.** $\sum_{k=1}^{\infty}\frac{1}{k(k+1)} = 1.$

*Beweis.* Durch Nachrechnen sieht man $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$ und damit konvergiert die Partialsummenfolge $\left(\sum_{k=1}^{n}\frac{1}{k(k+1)}\right)_{n\in\mathbb{N}} = \left(1 - \frac{1}{n+1}\right)_{n\in\mathbb{N}}$ gegen 1. □

**Satz 6.1.8** (Leibnizkriterium)**.** *Es sei* $x_k \geq 0$ *und* $(x_k)_{k\in\mathbb{N}}$ *eine monoton fallende Folge mit* $0 = \lim_{k\in\mathbb{N}} x_k$*, dann gilt:*

*a) Die alternierende Reihe* $\left(\sum_{k=1}^{n}(-1)^k x_k\right)_{n\in\mathbb{N}}$ *ist konvergent.*

*b) Für die Partialsummenfolge* $s_n := \sum_{k=1}^{n}(-1)^k x_k$ *gilt:*

$$s_{2n+1} \leq \sum_{k=1}^{\infty}(-1)^k x_k \quad \leq \quad s_{2n} \tag{6.1.4}$$

$$\left|\sum_{k=1}^{\infty}(-1)^k x_k - s_n\right| \quad \leq \quad x_{n+1} \tag{6.1.5}$$

*Beweis.* Nach der Bildungsvorschrift für die alternierende Reihe, da $x_k \geq 0$ und die Folge $(x_k)_{k\in\mathbb{N}}$ monoton fallend ist, gilt:

$$\begin{aligned}
s_{2n+1} &= s_{2n} - x_{2n+1} \leq s_{2n} \\
s_{2n+2} &= s_{2n} - (x_{2n+1} - x_{2n+2}) \leq s_{2n} \\
s_{2n+3} &= s_{2n+1} + (x_{2n+2} - x_{2n+3}) \geq s_{2n+1}
\end{aligned}$$

Daher ist die Teilfolge $(s_{2n})_{n\in\mathbb{N}}$ monoton fallend und die Teilfolge $(s_{2n+1})_{n\in\mathbb{N}}$ monoton steigend. Ferner ist $s_{2n} \leq s_2$, also wegen $s_{2n+1} \leq s_{2n}$ die Teilfolge $(s_{2n+1})_{n\in\mathbb{N}}$ nach oben durch $s_2$ beschränkt. Ebenso ist $s_1 \leq s_{2n+1} \leq s_{2n}$ für alle $n \in \mathbb{N}$ und damit ist $s_1$ untere Schranke von $(s_{2n})_{n\in\mathbb{N}}$. Damit ist $(s_{2n})_{n\in\mathbb{N}}$ monoton fallend und nach unten beschränkt, also nach Korollar 5.2.5 konvergent mit

$$\Sigma := \lim_{n\to\infty} s_{2n} = \inf\{s_{2n} : n \in \mathbb{N}\}.$$

Ebenso ist $(s_{2n+1})_{n\in\mathbb{N}}$ monoton steigend und nach oben beschränkt, also konvergent mit

$$\sigma := \lim_{n\to\infty} s_{2n+1} = \sup\{s_{2n+1} : n \in \mathbb{N}\}.$$

Wegen $s_{2n} - s_{2n+1} = x_{2n+1} \overset{n\to\infty}{\longrightarrow} 0$ ist $\Sigma - \sigma = 0$ und damit gilt $\Sigma = \sum_{k=1}^{\infty}(-1)^k x_k = \sigma$ analog zu Aufgabe 3 Tutoriumsblatt 9 [1]. Insbesondere gilt daher

$$s_{2n+1} \leq \sup\{s_{2n+1} : n \in \mathbb{N}\} = \sum_{k=1}^{\infty}(-1)^k x_k = \inf\{s_{2n} : n \in \mathbb{N}\} \leq s_{2n}$$

für alle $n \in \mathbb{N}$, was (6.1.4) zeigt und daraus folgt wegen $\left|\sum_{k=1}^{\infty}(-1)^k x_k - s_{2n}\right| \leq s_{2n} - s_{2n+1}$ und $\left|\sum_{k=1}^{\infty}(-1)^k x_k - s_{2n+1}\right| \leq s_{2n} - s_{2n+1}$ die Abschätzung (6.1.5). □

## 6.2 Absolut konvergente Reihen

**Definition 6.2.1.** *Es sei $(V, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum.*

- *Eine Reihe $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ in $V$ heißt* **absolut konvergent**, *wenn die Reihe $\left(\sum_{k=1}^{n} \|x_k\|\right)_{n\in\mathbb{N}}$ in $\mathbb{R}$ konvergiert.*
- *Ist $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ eine Reihe in $V$ und $\left(\sum_{k=1}^{n} y_k\right)_{n\in\mathbb{N}}$ eine Reihe in $\mathbb{R}$ mit $y_k \geq 0$ und $\|x_k\| \leq y_k$ für alle $k \in \mathbb{N}$, dann heißt $\left(\sum_{k=1}^{n} y_k\right)_{n\in\mathbb{N}}$ eine* **Majorante** *von $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$.*

**Satz 6.2.2** (Majorantenkriterium)**.** *Es sei $(X, \|\cdot\|)$ ein $\mathbb{K}$-Banachraum, $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ eine Reihe in $X$ und $\left(\sum_{k=1}^{n} y_k\right)_{n\in\mathbb{N}}$ eine konvergente Majorante. Dann konvergieren auch*

[1] dort hatten wir im Spezialfall gezeigt, daß es für eine Folge $(a_n)_{n\in\mathbb{N}}$ in einem metrischen Raum $(X, d)$ für die $(a_{2n})_{n\in\mathbb{N}}$ und $(a_{2n+1})_{n\in\mathbb{N}}$ gegen denselben Grenzwert $a \in X$ konvergieren, für jedes $\varepsilon > 0$ natürliche Zahlen $M(\varepsilon), N(\varepsilon) \in \mathbb{N}$ existieren mit $d(a, a_{2n}) < \varepsilon$ für alle $2n \geq M(\varepsilon)$ und $d(a, a_{2n+1}) < \varepsilon$ für alle $2n+1 \geq N(\varepsilon)$. Damit ist $d(a, a_m) < \varepsilon$ für alle $m \geq \max\{N(\varepsilon), M(\varepsilon)\}$ und somit ist $(a_n)_{n\in\mathbb{N}}$ ebenfalls konvergiert mit $a = \lim_{n\to\infty} a_n$.

$\left(\sum_{k=1}^{n} \|x_k\|\right)_{n\in\mathbb{N}}$ *und* $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ *und es gilt:*

$$\left\|\sum_{k=1}^{\infty} x_k\right\| \leq \sum_{k=1}^{\infty} y_k \qquad \text{und} \qquad \sum_{k=1}^{\infty} \|x_k\| \leq \sum_{k=1}^{\infty} y_k. \tag{6.2.1}$$

*Beweis.* Da die Reihe $\left(\sum_{k=1}^{n} y_k\right)_{n\in\mathbb{N}}$ konvergiert, ist sie eine Cauchyfolge und deshalb gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so daß $\sum_{k=m+1}^{n} y_k < \varepsilon$ für alle $m, n \geq N$ gilt. Da $\|x_k\| \leq y_k$ ist, folgt aus der Dreiecksungleichung

$$\left\|\sum_{k=m+1}^{n} x_k\right\| \leq \sum_{k=m+1}^{n} \|x_k\| \leq \sum_{k=m+1}^{n} y_k < \varepsilon$$

für $n, m \geq N$. Somit sind $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ und $\left(\sum_{k=1}^{n} \|x_k\|\right)_{n\in\mathbb{N}}$ Cauchyfolgen, also konvergiert die Reihe $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$, da $X$ ein Banachraum ist und $\left(\sum_{k=1}^{n} \|x_k\|\right)_{n\in\mathbb{N}}$ konvergiert in $\mathbb{R}$, da $\mathbb{R}$ nach Satz 5.1.10 vollständig ist. Für $x := \sum_{k=1}^{\infty} x_k$ gibt es für jedes $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ mit $\left\|x - \sum_{k=1}^{n} x_k\right\| < \varepsilon$ für jedes $n \geq N(\varepsilon)$, also folgt $\left|\, \|x\| - \left\|\sum_{k=1}^{n} x_k\right\| \,\right| \leq \left\|x - \sum_{k=1}^{n} x_k\right\| < \varepsilon$ für alle $n \geq N(\varepsilon)$ nach der umgekehrten Dreiecksungleichung, dh.

$$\|x\| = \left\|\sum_{k=1}^{\infty} x_k\right\| = \lim_{n\to\infty} \left\|\sum_{k=1}^{n} x_k\right\| \tag{6.2.2}$$

Nach Voraussetzung gilt $\left\|\sum_{k=1}^{n} x_k\right\| \leq \sum_{k=1}^{n} \|x_k\| \leq \sum_{k=1}^{n} y_k$ und daher folgt

$$\left\|\sum_{k=1}^{\infty} x_k\right\| = \lim_{n\to\infty} \left\|\sum_{k=1}^{n} x_k\right\| \leq \lim_{n\to\infty} \sum_{k=1}^{n} \|x_k\| = \sum_{=1}^{\infty} \|x_k\| \leq \lim_{n\to\infty} \sum_{k=1}^{n} y_k = \sum_{k=1}^{\infty} y_k$$

nach (6.2.2) und Lemma 5.2.2. □

**Korollar 6.2.3.** *Ist $X$ ein Banachraum, so ist jede absolut konvergente Reihe in $X$ auch konvergent.*

**Beispiel 6.2.4.**

- Wegen $\frac{1}{(k+1)^2} \le \frac{1}{k(k+1)}$ hat die Reihe $\left(\sum_{k=2}^{n} \frac{1}{k^2}\right)_{n\in\mathbb{N}}$ die konvergente Majorante $\left(\sum_{k=1}^{n} \frac{1}{k(k+1)}\right)_{n\in\mathbb{N}}$, vgl. Beispiel 6.1.7, daher existiert auch $\sum_{k=1}^{\infty} \frac{1}{k^2} = 1 + \sum_{k=2}^{\infty} \frac{1}{k^2}$ und es gilt $\sum_{k=1}^{\infty} \frac{1}{k^2} \le 2$.

- Für alle $j \ge 2$ und $k \in \mathbb{N}$ ist $\frac{1}{k^j} \le \frac{1}{k^2}$, daher ist nach dem Majorantenkriterium und dem letzten Beispiel auch $\left(\sum_{k=1}^{n} \frac{1}{k^j}\right)_{n\in\mathbb{N}}$ konvergent und es gilt $\sum_{k=1}^{\infty} \frac{1}{k^j} \le 2$.

**Satz 6.2.5** (Wurzelkriterium). *Es sei $(X, \|\cdot\|)$ ein $\mathbb{K}$-Banachraum, es sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $X$ und*

$$L := \overline{\lim_{n\to\infty}} \sqrt[n]{\|x_n\|}.$$

*Dann gilt:*

*a) Ist $L < 1$, dann ist die Reihe $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ absolut konvergent.*

*b) Ist $L > 1$, dann ist die Reihe $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ nicht konvergent.*

*Beweis.*

- Im Fall $L := \overline{\lim_{n\to\infty}} \sqrt[n]{\|x_n\|} = \limsup_{n\to\infty} \sqrt[n]{\|x_n\|} < 1$ wähle $q \in \mathbb{R}$ mit $L < q < 1$, dann gibt es nach Lemma 5.5.4 ein $N \in \mathbb{N}$ mit $\sqrt[n]{\|x_n\|} < q$ für alle $n \ge N$. Daher hat $\left(\sum_{k=N}^{n} \|x_k\|\right)_{n\ge N}$ die konvergente Majorante $\left(\sum_{k=N}^{n} q^k\right)_{n\ge N}$ und daher ist $\left(\sum_{k=1}^{n} \|x_k\|\right)_{n\in\mathbb{N}}$ konvergent.

- Im Fall $L := \overline{\lim_{n\to\infty}} \sqrt[n]{\|x_n\|} > 1$ gibt es eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\in\mathbb{N}}$ mit $\|x_{n_k}\| \ge 1$, also ist $(x_n)_{n\in\mathbb{N}}$ keine Nullfolge und damit $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ nicht konvergent.

$\square$

**Satz 6.2.6** (Quotientenkriterium)**.** *Es sei $(X, \|\cdot\|)$ ein $\mathbb{K}$-Banachraum, $(x_n)_{n\in\mathbb{N}}$ eine Folge in $X$ mit $x_k \neq \mathbf{0}$ für alle $k \in \mathbb{N}$. Dann gilt:*

*a) Ist $\displaystyle\varlimsup_{n\to\infty} \frac{\|x_{n+1}\|}{\|x_n\|} < 1$, so ist die Reihe $\displaystyle\left(\sum_{k=1}^n x_k\right)_{n\in\mathbb{N}}$ absolut konvergent.*

*b) Ist $\displaystyle\varliminf_{n\to\infty} \frac{\|x_{n+1}\|}{\|x_n\|} > 1$, so konvergiert die Reihe $\displaystyle\left(\sum_{k=1}^n x_k\right)_{n\in\mathbb{N}}$ nicht.*

*Beweis.*

- Es sei $q := \displaystyle\varlimsup_{n\to\infty} \frac{\|x_{n+1}\|}{\|x_n\|} < 1$, dann wähle $p \in \mathbb{R}$ mit $q < p < 1$ und dann gibt es nach Lemma 5.5.4 ein $N \in \mathbb{N}$ mit $\frac{\|x_{n+1}\|}{\|x_n\|} \leq p$ für alle $n \geq N$. Daher gilt für alle $k \geq N$

$$\|x_k\| \leq p\|x_{k-1}\| \leq p^2\|x_{k-2}\| \leq ... \leq p^{k-N}\|x_N\|$$

und daher ist $\displaystyle\left(\|x_N\| \sum_{k=N}^n p^{k-N}\right)_{n\geq N}$ eine konvergente Majorante von $\displaystyle\left(\sum_{k=N}^n \|x_k\|\right)_{n\geq N}$.

Nach Majorantenkriterium konvergiert auch $\displaystyle\left(\sum_{k=1}^n \|x_k\|\right)_{n\in\mathbb{N}}$.

- Es sei $q := \displaystyle\varliminf_{n\to\infty} \frac{\|x_{n+1}\|}{\|x_n\|} > 1$, dann gibt es für $p \in ]1, q[$ ein $N \in \mathbb{N}$ mit $\|x_{k+1}\| > p\|x_k\|$ für alle $k \geq N$. Damit ist $(x_k)_{k\in\mathbb{N}}$ keine Nullfolge, also $\displaystyle\left(\sum_{k=1}^n x_k\right)_{n\in\mathbb{N}}$ nach Bemerkung 6.1.4 nicht konvergent. □

**Bemerkung 6.2.7.** $\displaystyle\lim_{n\to\infty} \frac{n^2}{(n+1)^2} = \lim_{n\to\infty} \frac{n}{n+1} = 1$ und nach Satz 5.3.5 und den Rechenregeln für konvergente Folgen folgt $\displaystyle\limsup_{n\to\infty} \sqrt[n]{n^2} = \lim_{n\to\infty} \sqrt[n]{n}^2 = \lim_{n\to\infty} \sqrt[n]{n} = 1$. Damit machen weder Quotientenkriterium noch Wurzelkriterium eine Aussage über die Konvergenz bzw. Nichtkonvergenz der Reihen $\displaystyle\left(\sum_{k=1}^n \frac{1}{k^2}\right)_{n\in\mathbb{N}}$ und $\displaystyle\left(\sum_{k=1}^n \frac{1}{k}\right)_{n\in\mathbb{N}}$. Diese beiden Beispiele zeigen, daß in den Fällen, in denen die Voraussetzungen der Konvergenzkriterien nicht erfüllt sind, sowohl Konvergenz als auch Nichtkonvergenz der Reihe auftreten kann.

**Satz 6.2.8** (Verdichtungskriterium)**.** *Es sei $x_n \geq 0$, $(x_n)_{n\in\mathbb{N}}$ monoton fallend und für die streng monoton steigende Abbildung $f : \mathbb{N} \to \mathbb{N}$ existieren $\alpha, \beta \in \mathbb{R}$ mit*

$$1 < \alpha \leq \inf\left\{\frac{f(n+1)}{f(n)} : n \in \mathbb{N}\right\} \leq \sup\left\{\frac{f(n+1)}{f(n)} : n \in \mathbb{N}\right\} \leq \beta.$$

*Dann sind äquivalent:*

*a) $\displaystyle\left(\sum_{k=1}^n x_k\right)_{n\in\mathbb{N}}$ ist konvergent.*

*b)* $\left(\sum_{k=1}^{n} f(k)x_{f(k)}\right)_{n\in\mathbb{N}}$ *ist konvergent.*

*Beweis.* Für $k \geq 2$ ist $\alpha \leq \dfrac{f(k)}{f(k-1)}$ oder $0 \leq f(k) - \alpha f(k-1)$ und nach Addition von $(\alpha-1)f(k)$ gilt $(\alpha-1)f(k) \leq \alpha f(k) - \alpha f(k-1)$ also folgt mit $\alpha > 1$:

$$f(k) \leq \frac{\alpha}{\alpha-1}(f(k) - f(k-1)) \qquad \text{für } k \geq 2. \tag{6.2.3}$$

Setzt man noch $f(0) := 0$, so gilt diese Abschätzung auch noch für $k = 1$.
Aus $\dfrac{f(k+1)}{f(k)} \leq \beta$ folgt $f(k+1) \leq \beta f(k)$ und nach Subtraktion von $f(k)$ folgt

$$f(k+1) - f(k) \leq (\beta-1)f(k). \tag{6.2.4}$$

„$\Rightarrow$“ Ist $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ konvergent, so gilt

$$\begin{aligned}\sum_{k=1}^{n} f(k)x_{f(k)} &\leq \frac{\alpha}{\alpha-1}\sum_{k=1}^{n}(f(k)-f(k-1))x_{f(k)} \leq \frac{\alpha}{\alpha-1}\sum_{k=1}^{n}\sum_{l=f(k-1)+1}^{f(k)} x_l \\ &= \frac{\alpha}{\alpha-1}\sum_{k=1}^{f(n)} x_k\end{aligned}$$

daher hat $\left(\sum_{k=1}^{n} f(k)x_{f(k)}\right)_{n\in\mathbb{N}}$ die konvergente Majorante $\left(\dfrac{\alpha}{\alpha-1}\sum_{k=1}^{f(n)} x_k\right)_{n\in\mathbb{N}}$ ist also konvergent.

„$\Leftarrow$“ Es sei $\left(\sum_{k=1}^{n} f(k)x_{f(k)}\right)_{n\in\mathbb{N}}$ konvergent. Da $f : \mathbb{N} \to \mathbb{N}$ streng monoton steigend ist, gilt $n \leq f(n)$ und weil $x_k \geq 0$ monoton fallend ist, folgt:

$$\begin{aligned}\sum_{k=1}^{n} x_k &\leq \sum_{k=1}^{f(n)} x_k \leq \sum_{k=1}^{f(1)-1} x_k + \sum_{k=1}^{n}\sum_{l=f(k)}^{f(k+1)-1} x_l \leq \sum_{k=1}^{f(1)-1} x_k + \sum_{k=1}^{n}(f(k+1)-f(k))x_{f(k)} \\ &\leq \sum_{k=1}^{f(1)-1} x_k + (\beta-1)\sum_{k=1}^{n} f(k)x_{f(k)}\end{aligned}$$

und damit hat $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ eine konvergente Majorante. $\square$

**Bemerkung 6.2.9.**

- Das bekannteste Beispiel für eine Funktion $f : \mathbb{N} \to \mathbb{N}$, die die Bedingungen im Verdichtungskriterium erfüllt, ist $\begin{array}{rcl} f : \mathbb{N} & \to & \mathbb{N} \\ n & \mapsto & 2^n \end{array}$.
- Mit dem Verdichtungskriterium kann man leicht zeigen, daß für $\alpha > 1$ die Reihe $\left(\sum_{k=1}^{n} k^{-\alpha}\right)_{n\in\mathbb{N}}$ absolut konvergent ist, (beachte: Wir haben bisher $k^\alpha$ nur für $\alpha \in \mathbb{Q}$ definiert!), denn

$$\sum_{k=1}^{n} 2^k (2^k)^{-\alpha} = \sum_{k=1}^{n} (2^k)^{1-\alpha} = \sum_{k=1}^{n} (2^{1-\alpha})^k$$

konvergiert wegen $|2^{1-\alpha}| < 1$ als Beispiel einer geometrischen Reihe.

**Definition 6.2.10.** *Es sei $(X, \|\cdot\|)$ ein $\mathbb{K}$-Banachraum, $\varphi : \mathbb{N} \to \mathbb{N}$ bijektiv und $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ eine Reihe in $X$, dann heißt $\left(\sum_{k=1}^{n} x_{\varphi(k)}\right)_{n\in\mathbb{N}}$ die mit $\varphi$* **umgeordnete Reihe**.

**Satz 6.2.11.** *Es sei $(X, \|\cdot\|)$ ein $\mathbb{K}$-Banachraum und $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ eine absolut konvergente Reihe in $X$ mit Grenzwert $x = \sum_{k=1}^{\infty} x_k$. Dann sind für jedes bijektive $\varphi : \mathbb{N} \to \mathbb{N}$ auch $\left(\sum_{k=1}^{n} x_{\varphi(k)}\right)_{n\in\mathbb{N}}$ und $\left(\sum_{k=1}^{n} \|x_{\varphi(k)}\|\right)_{n\in\mathbb{N}}$ konvergent und es gilt:*

$$\sum_{k=1}^{\infty} x_k = \sum_{k=1}^{\infty} x_{\varphi(k)} \qquad \textit{und} \qquad \sum_{k=1}^{\infty} \|x_k\| = \sum_{k=1}^{\infty} \|x_{\varphi(k)}\|. \tag{6.2.5}$$

*Beweis.* Wegen absoluter Konvergenz gibt es für jedes $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ mit $\sum_{k=l}^{m} \|x_k\| < \frac{\varepsilon}{3}$ für alle $m, l \geq N(\varepsilon)$, denn $\left(\sum_{k=1}^{n} \|x_k\|\right)_{n\in\mathbb{N}}$ ist eine Cauchyfolge. Wähle nun $P(\varepsilon) := \max\{\varphi^{-1}(1), \varphi^{-1}(2), ..., \varphi^{-1}(N(\varepsilon))\}$, dann gilt

$$\left\|\sum_{k=1}^{n} x_k - \sum_{k=1}^{n} x_{\varphi(k)}\right\| \leq 2 \sum_{k=N(\varepsilon)+1}^{n} \|x_k\| < \frac{2\varepsilon}{3}$$

für alle $n \geq \max\{P(\varepsilon), N(\varepsilon)\}$, denn die Reihenglieder $x_1 = x_{\varphi(\varphi^{-1}(1))}, ..., x_{N(\varepsilon)} = x_{\varphi(\varphi^{-1}(N(\varepsilon)))}$ kommen in den beiden Summen vor und heben sich damit weg. Da die absolut konvergente Reihe $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ nach Korollar 6.2.3 auch konvergiert gibt es ein $M(\varepsilon) \in \mathbb{N}$ mit

$\left\| \sum_{k=1}^{\infty} x_k - \sum_{k=1}^{n} x_k \right\| < \frac{\varepsilon}{3}$ für alle $n \geq M(\varepsilon)$. Für $n \geq \max\{M(\varepsilon), N(\varepsilon), P(\varepsilon)\}$ gilt also

$$\left\| \sum_{k=1}^{\infty} x_k - \sum_{k=1}^{n} x_{\varphi(k)} \right\| \leq \left\| \sum_{k=1}^{\infty} x_k - \sum_{k=1}^{n} x_k \right\| + \left\| \sum_{k=1}^{n} x_k - \sum_{k=1}^{n} x_{\varphi(k)} \right\| < \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

was die Konvergenz von $\left( \sum_{k=1}^{n} x_{\varphi(k)} \right)_{n\in\mathbb{N}}$ und die Gleichheit $\sum_{k=1}^{\infty} x_k = \sum_{k=1}^{\infty} x_{\varphi(k)}$ zeigt. □

**Satz 6.2.12** (Riemannscher Umordnungssatz)**.** *Es sei* $\left( \sum_{k=1}^{n} x_k \right)_{n\in\mathbb{N}}$ *eine konvergente Reihe in* $\mathbb{R}$*, die nicht absolut konvergent ist. Dann gibt es für jedes* $c \in \widehat{\mathbb{R}}$ *eine bijektive Abbildung* $\varphi_c : \mathbb{N} \longrightarrow \mathbb{N}$*, so daß für die umgeordnete Reihe* $\left( \sum_{k=1}^{n} x_{\varphi_c(k)} \right)_{n\in\mathbb{N}}$ *gilt:*

$$c = \lim_{n\to\infty} \sum_{k=1}^{n} x_{\varphi_c(k)}$$

**Beispiel 6.2.13.** Ein Beispiel, das die Problematik beim Umordnen erläutert – und „nebenbei“ auch eine Idee gibt, wie der Beweis des Riemannschen Umordnungssatzes geht – ist die alternierende harmonische Reihe $\left( \sum_{k=1}^{n} (-1)^k \frac{1}{k} \right)_{n\in\mathbb{N}}$, die nach dem Leibnizkriterium konvergiert, die aber nicht absolut konvergiert (harmonische Reihe!). Die beiden Folgen $\left( \frac{1}{2k} \right)_{k\in\mathbb{N}}$ und $\left( \frac{1}{2k+1} \right)_{k\in\mathbb{N}}$ sind monoton fallend, $\lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{2k} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{2k+1} = \infty$ in $\widehat{\mathbb{R}}$. Daher findet man zu vorgegebenem $c \in [0, \infty[$ ein $n_1 \in \mathbb{N}$ mit $\sum_{k=1}^{n_1-1} \frac{1}{2k} < c \leq \sum_{k=1}^{n_1} \frac{1}{2k}$ und dann ein $m_1 \in \mathbb{N}$ mit

$$\sum_{k=1}^{n_1} \frac{1}{2k} - \sum_{k=0}^{m_1-1} \frac{1}{2k+1} \geq c > \sum_{k=1}^{n_1} \frac{1}{2k} - \sum_{k=0}^{m_1} \frac{1}{2k+1}$$

anschließend ein $n_2 > n_1$ mit

$$\sum_{k=1}^{n_1} \frac{1}{2k} - \sum_{k=0}^{m_1-1} \frac{1}{2k+1} \sum_{k=n_1+1}^{n_2-1} \frac{1}{2k} \leq c < \sum_{k=1}^{n_1} \frac{1}{2k} - \sum_{k=0}^{m_1} \frac{1}{2k+1} + \sum_{k=n_1+1}^{n_2} \frac{1}{2k}$$

usw. dann konvergiert die umgeordnete Reihe gegen $c$. Die beiden folgenden Bilder deuten an, wie eine Umordnung für einen Grenzwert $c > 0$ bzw. für $c = \infty$ aussieht:

## 6.3 Potenzreihen

**Definition 6.3.1.** *Es sei $(X, \|\cdot\|)$ ein $\mathbb{C}$-Banachraum, $a \in \mathbb{C}$ und $(c_n)_{n\in\mathbb{N}_0}$ eine Folge in $X$. Dann heißt*

$$\left(\sum_{n=0}^{N} c_n(z-a)^n\right)_{N\in\mathbb{N}_0}$$

**Potenzreihe** *zum* **Entwicklungspunkt** *$a$. Es sei*

$$L := \overline{\lim_{n\to\infty}} \sqrt[n]{\|c_n\|},$$

*dann heißt*

$$\rho := \begin{cases} 0 & \text{falls } L = \infty \\ \frac{1}{L} & \text{falls } L \in ]0,\infty[ \\ \infty & \text{falls } L = 0 \end{cases} \tag{6.3.1}$$

*der* **Konvergenzradius** *und*

$$K(a,\rho) := \{z \in \mathbb{C} : |z-a| < \rho\} \tag{6.3.2}$$

*das* **Innere des Konvergenzkreises** *um den* **Entwicklungspunkt** *$a$.*

**Satz 6.3.2.** *Es sei $(X, \|\cdot\|)$ ein $\mathbb{C}$-Banachraum, $a \in \mathbb{C}$ und $(c_n)_{n\in\mathbb{N}_0}$ eine Folge in $X$ und $\rho$ der Konvergenzradius von $\left(\sum_{n=0}^{N} c_n(z-a)^n\right)_{N\in\mathbb{N}_0}$.*

- *Ist $\rho = 0$, so konvergiert die Reihe $\left(\sum_{n=0}^{N} c_n(z-a)^n\right)_{N\in\mathbb{N}_0}$ für kein $z \in \mathbb{C}\backslash\{a\}$ und degeneriert für $z = a$ zu $c_0$.*
- *Ist $\rho \in ]0,\infty[$, dann ist die Reihe $\left(\sum_{n=0}^{N} c_n(z-a)^n\right)_{N\in\mathbb{N}_0}$ für alle $z \in \mathbb{C}$ mit $|z-a| < \rho$ absolut konvergent und divergent für alle $z \in \mathbb{C}$ mit $|z-a| > \rho$.*
- *Ist $\rho = \infty$, so konvergiert $\left(\sum_{n=0}^{N} c_n(z-a)^n\right)_{N\in\mathbb{N}_0}$ für alle $z \in \mathbb{C}$ absolut.*

*Beweis.* Es sei $L := \overline{\lim_{n\to\infty}} \sqrt[n]{\|c_n\|}$.

- Im Fall $L \in ]0,\infty[$ sei $z \neq a$, dann ist

$$\limsup_{n\to\infty} \sqrt[n]{\|c_n\|\ |z-a|^n} = |z-a| \limsup_{n\to\infty} \sqrt[n]{\|c_n\|} = |z-a|L.$$

  In diesem Fall folgt absolute Konvergenz von $\left(\sum_{n=0}^{N} c_n(z-a)^n\right)_{N\in\mathbb{N}_0}$ nach dem Wurzelkriterium falls $|z-a| < \frac{1}{L}$ ist. Aus $|z-a| > \frac{1}{L}$ folgt mit dem Wurzelkriterium die Divergenz der Reihe.

- Ist $L = 0$, so ist
$$\limsup_{n\to\infty} \sqrt[n]{\|c_n\|\ |z-a|^n} = |z-a| \limsup_{n\to\infty} \sqrt[n]{\|c_n\|} = |z-a|L = 0.$$
für jedes $z \in \mathbb{C}$ und daher konvergiert $\left(\sum_{n=0}^{N} c_n(z-a)^n\right)_{N\in\mathbb{N}_0}$ nach dem Wurzelkriterium für alle $z \in \mathbb{C}$ absolut.
- Im Fall $L = \infty$ sei $z \in \mathbb{C}\backslash\{a\}$ und $r := \frac{|z-a|}{2} > 0$; dann gibt es eine Teilfolge, so daß $\lim\limits_{j\to\infty} \sqrt[n_j]{\|c_{n_j}\|} \geq r$ ist. Insbesondere ist dann
$$\overline{\lim_{n\to\infty}} \sqrt[n]{\|c_n\|\ |z-a|^n} \geq \limsup_{j\to\infty} \sqrt[n_j]{\|c_{n_j}\|\ |z-a|^{n_j}} \geq r|z-a| > 1$$
also $\left(\sum_{n=0}^{N} c_n(z-a)^n\right)_{N\in\mathbb{N}_0}$ nach dem Wurzelkriterium nicht konvergent.

□

**Satz 6.3.3.** *Es sei $(X, \|\cdot\|)$ ein $\mathbb{C}$-Banachraum, $(c_n)_{n\in\mathbb{N}_0}$ eine Folge in $X$ mit $c_n \neq \mathbf{0}$ und es existiere*
$$q := \lim_{n\to\infty} \frac{\|c_{n+1}\|}{\|c_n\|} \in [0,\infty], \tag{6.3.3}$$
*dann ist* $\rho := \left\{\begin{array}{ll} 0 & \textit{falls } q = \infty \\ \frac{1}{q} & \textit{falls } q \in ]0,\infty[ \\ \infty & \textit{falls } q = 0 \end{array}\right\}$ *der Konvergenzradius von* $\left(\sum_{n=0}^{N} c_n(z-a)^n\right)_{N\in\mathbb{N}_0}$.

*Beweis.* Unter Verwendung des Quotientenkriteriums erhällt man absolute Konvergenz der Reihe $\left(\sum_{n=0}^{N} c_n(z-a)^n\right)_{N\in\mathbb{N}_0}$ für alle $z \in \mathbb{C}$ mit $|z-a| < \rho$ in Fall $\rho \in ]0,\infty[$ bzw. für alle $z \in \mathbb{C}$ im Fall $\rho = \infty$ und keine Konvergenz der Reihe für alle $z \in \mathbb{C}$ mit $|z-a| > \rho$ in Fall $\rho \in ]0,\infty[$ bzw. für alle $z \in \mathbb{C}\backslash\{a\}$ im Fall $\rho = 0$. Die Kreisscheibe $K(a,\rho)$ ist damit gerade der Konvergenzkreis und $\rho$ der Konvergenzradius von $\left(\sum_{n=0}^{N} c_n(z-a)^n\right)_{N\in\mathbb{N}_0}$. □

**Bemerkung 6.3.4.** Ist $\rho \in ]0,\infty[$, so bekommt man auf dem Rand $\{z \in \mathbb{C} : |z-a| = \rho\}$ des Konvergenzkreises nur durch Betrachtung des Konvergenzradiuses $\rho$ keine Aussage über Konvergenz bzw. Nichtkonvergenz der Potenzreihe:

a) Bei $\left(\sum_{n=1}^{N} \frac{(-1)^n}{n} z^n\right)_{N\in\mathbb{N}}$ ist $\limsup\limits_{n\to\infty} \sqrt[n]{\frac{1}{n}} = 1$ – das folgt mit $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$ – oder $\lim\limits_{n\to\infty} \frac{n}{n+1} = 1$, also ist der Konvergenzradius $\rho = 1$ und damit $K(0,1) = \{z \in \mathbb{C} :$

$|z| < 1\}$ der Konvergenzkreis. Auf dem Rand $\{z \in \mathbb{C} : |z| = 1\}$ des Konvergenzkreises bekommt man mit anderen Konvergenzkriterien beispielsweise: Für $z = 1$ ist die Reihe $\left(\sum_{n=1}^{N}(-1)^n \frac{1}{n}\right)_{N\in\mathbb{N}}$ nach dem Leibnizkriterium konvergent, aber nicht absolut konvergent – harmonische Reihe. In $z = -1$ wird aus obiger Potenzreihe die nicht konvergente harmonische Reihe $\left(\sum_{n=1}^{N} \frac{1}{n}\right)_{N\in\mathbb{N}}$.

b) Auch für $\left(\sum_{n=0}^{N} z^n\right)_{N\in\mathbb{N}_0}$ ist der Konvergenzradius $\rho = 1$ und für $|z| = 1$ konvergiert $\left(\sum_{n=0}^{N} z^n\right)_{N\in\mathbb{N}_0}$ nicht, da die Reihenglieder $(z^n)_{n\in\mathbb{N}}$ keine Nullfolge bilden.

**Lemma 6.3.5.** *Es sei $X$ ein $\mathbb{C}-$Banachraum, $(b_n)_{n\in\mathbb{N}_0}$ eine Folge in $X$, $a \in \mathbb{C}$ und $\left(\sum_{n=0}^{N} b_n(z-a)^n\right)_{N\in\mathbb{N}_0}$ eine Potenzreihe mit Konvergenzradius $\rho > 0$. Dann gibt es für jedes $s \in ]0, \rho[$ ein $C_s \in ]0, \infty[$ mit*

$$\sup\{\|b_n\||z-a|^n : n \in \mathbb{N}_0, |z-a| \leq s\} \leq C_s. \tag{6.3.4}$$

*Beweis.* Da der Konvergenzradius $\rho > 0$ ist, gilt $L := \limsup_{n\to\infty} \sqrt[n]{\|b_n\|} \in [0, \infty[$.

- 1. Fall: $L \in ]0, \infty[$
  Nach Lemma 5.5.4 gibt es für jedes $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ mit $\sqrt[n]{\|b_n\|} \leq L + \varepsilon$ für alle $n \geq N(\varepsilon)$. Wähle $\varepsilon > 0$ mit $s = \frac{1}{L+\varepsilon}$, dann ist

  $$|z-a| \leq s = \frac{1}{L+\varepsilon} \leq \frac{1}{\sqrt[n]{\|b_n\|}}$$

  für alle $n \geq N(\varepsilon)$ oder

  $$\|b_n\|\,|z-a|^n \leq 1 \quad \text{für } n \geq N(\varepsilon) \tag{6.3.5}$$

- 2. Fall: $L = 0$
  Da $\sqrt[n]{\|b_n\|} \geq 0$, also $\liminf_{n\to\infty} \sqrt[n]{\|b_n\|} \geq 0$ ist, gilt

  $$\limsup_{n\to\infty} \sqrt[n]{\|b_n\|} = \lim_{n\to\infty} \sqrt[n]{\|b_n\|} = 0.$$

  Deshalb gibt es zu $|z-a| \in ]0, \infty[$ ein $N \in \mathbb{N}$ mit $\sqrt[n]{\|b_n\|} \leq \frac{1}{|z-a|}$ für $n \geq N$ oder

  $$\|b_n\|\,|z-a|^n \leq 1 \quad \text{für } n \geq N. \tag{6.3.6}$$

Da $\{\|b_j\|\,|z-a|^j : j = 0, ..., N-1\}$ beschränkt ist, folgt die Behauptung. □

## 6.4 Netze und summierbare Familien

**Definition 6.4.1.** *Es sei $\emptyset \neq I$ eine* **gerichtete Menge***, dh. $(I, \leq)$ ist eine geordnete Menge, so daß es zu je zwei Elementen $i, j \in I$ eine obere Schranke $k \in I$ mit $i \leq k, j \leq k$ gibt. Es sei $(X, d)$ ein metrischer Raum und $I$ gerichtet, dann heißt eine* **Familie** *$(x_i)_{i \in I}$ von Elementen in $X$ – dh. eine Abbildung $\begin{array}{rccc} x : & I & \to & X \\ & i & \mapsto & x_i \end{array}$ – ein* **Netz** *in $X$. $a \in X$ heißt* **Limes** *oder* **Grenzwert** *des Netzes $(x_i)_{i \in I}$, wenn es für alle $\varepsilon > 0$ ein $j = j(\varepsilon) \in I$ gibt, so daß für alle $i \in I$ mit $i \geq j(\varepsilon)$ gilt:*

$$d(x, x_i) < \varepsilon. \tag{6.4.1}$$

*Schreibweise: $a = \lim_{i \in I} x_i$ oder $x_i \to a$.*

**Lemma 6.4.2.** *Ist $(X, d)$ ein metrischer Raum, so hat ein Netz $(x_i)_{i \in I}$ in $X$ höchstens einen Grenzwert.*

*Beweis.* Ist $(X, d)$ ein metrischer Raum, $(x_i)_{i \in I}$ ein Netz in $X$ und $a, b \in X$ mit $x_i \to a$ und $x_i \to b$. Dann gibt es für jedes $\varepsilon > 0$ ein

- $j_a(\varepsilon) \in I$ mit $d(x_i, a) < \varepsilon$ für alle $i \in I$ mit $i \geq j_a(\varepsilon)$.
- $j_b(\varepsilon) \in I$ mit $d(x_i, b) < \varepsilon$ für alle $i \in I$ mit $i \geq j_b(\varepsilon)$.

Da die Indexmenge $I$ gerichtet ist, gibt es $j \in I$ mit $j \geq j_a(\varepsilon)$ und $j \geq j_b(\varepsilon)$, also folgt $d(x_j, a) < \varepsilon$ und $d(x_j, b) < \varepsilon$. Mit der Dreiecksungleichung folgt dann

$$d(a, b) \leq d(a, x_j) + d(x_j, b) < 2\varepsilon. \tag{6.4.2}$$

Weil dies für alle $\varepsilon > 0$ gilt, folgt daraus $d(a, b) = 0$, also $a = b$. □

**Definition 6.4.3.** *Ist $(X, d)$ ein metrischer Raum, $(x_i)_{i \in I}$ ein Netz in $X$, so heißt $(x_i)_{i \in I}$ ein* **Cauchynetz***, wenn es für alle $\varepsilon > 0$ ein $j = j(\varepsilon) \in I$ gibt, so daß für alle $k, l \in I$ mit $k \geq j(\varepsilon)$ und $l \geq j(\varepsilon)$ gilt:*

$$d(x_k, x_l) < \varepsilon. \tag{6.4.3}$$

**Bemerkung 6.4.4.** Im Fall $I = \mathbb{N}$ sind "Folge" und "Netz", "Cauchyfolge" und "Cauchynetz" und ebenso "Grenzwert von Folgen" und "Grenzwert von Netzen" identisch.

**Lemma 6.4.5.** *Es sei $(X, d)$ ein metrischer Raum und $(x_i)_{i \in I}$ ein konvergentes Netz in $X$. Dann ist $(x_i)_{i \in I}$ ein Cauchynetz.*

*Beweis.* Es sei $x := \lim_{i \in I} x_i \in X$ und $\varepsilon > 0$, dann gibt es $j = j(\varepsilon) \in I$, so daß für alle $i \in I$ mit $i \geq j(\varepsilon)$ gilt: $d(x, x_i) < \frac{\varepsilon}{2}$. Sind also $k, l \in I$ mit $k \geq j(\varepsilon)$ und $l \geq j(\varepsilon)$, dann ist

$$d(x_k, x_l) \leq d(x_k, x) + d(x, x_l) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

dh. $(x_i)_{i \in I}$ ist ein Cauchynetz. □

**Satz 6.4.6.** *Es sei $(X, d)$ ein vollständiger metrischer Raum und $(x_i)_{i\in I}$ ein Cauchynetz in $X$, dann konvergiert $(x_i)_{i\in I}$.*

*Beweis.* Ist $(x_i)_{i\in I}$ ein Cauchynetz in $X$, so gibt es für jedes $k \in \mathbb{N}$ ein $\varphi(k) \in I$, so daß $d(x_i, x_j) < \frac{1}{k}$ für alle $i, j \in I$ mit $i \geq \varphi(k)$ und $j \geq \varphi(k)$. Ohne Einschränkung darf man $\varphi(k+1) \geq \varphi(k)$ annehmen, denn da $I$ gerichtet ist, kann man gegebenenfalls $\varphi(k+1)$ durch eine obere Schranke von $\varphi(k)$ und $\varphi(k+1)$ ersetzen. Dann ist $(x_{\varphi(k)})_{k\in\mathbb{N}}$ eine Cauchyfolge in $X$; diese konvergiert nach Voraussetzung und so sei $x := \lim\limits_{k\to\infty} x_{\varphi(k)}$ der Grenzwert. Zum Nachweis von $x = \lim\limits_{i\in I} x_i$ sei $\varepsilon > 0$, dann existiert wegen $x = \lim\limits_{k\to\infty} x_{\varphi(k)}$ ein $N = N(\varepsilon) \in \mathbb{N}$ mit $d(x, x_{\varphi(k)}) < \frac{\varepsilon}{2}$ für alle $k \geq N(\varepsilon)$ und nach evtl. Vergrößern von $N(\varepsilon)$ darf man noch $\dfrac{1}{N(\varepsilon)} < \dfrac{\varepsilon}{2}$ annehmen. Ist nun $i \in I$ mit $i \geq \varphi(N(\varepsilon))$, so folgt aus der Dreiecksungleichung

$$d(x, x_i) \leq d(x, x_{\varphi(N(\varepsilon))}) + d(x_{\varphi(N(\varepsilon))}, x_i) < \frac{\varepsilon}{2} + \frac{1}{N(\varepsilon)} < \varepsilon,$$

dh. $x = \lim\limits_{i\in I} x_i$ ist der Grenzwert von $(x_i)_{i\in I}$. □

**Definition 6.4.7.** *Es sei $(V, \|\cdot\|)$ ein normierter Vektorraum, $\emptyset \neq I$ eine Indexmenge und $(x_i)_{i\in I}$ eine Familie in $V$. Zu jeder endlichen Teilmenge $H \in \mathcal{E}(I)$ sei*

$$s_H := \sum_{i\in H} x_i$$

*die* **Partialsumme** *von $(x_i)_{i\in I}$ zu $H$. Durch*

$$H \leq J :\Leftrightarrow H \subseteq J \tag{6.4.4}$$

*wird $\mathcal{E}(I)$ zu einer gerichteten Menge und damit*

$$(s_H)_{H\in\mathcal{E}(I)} = (\sum_{i\in H} x_i)_{H\in\mathcal{E}(I)}$$

*ein Netz in $V$. $(x_i)_{i\in I}$ heißt* **summierbare Familie***, wenn*

$$\lim_{H\in\mathcal{E}(I)} s_H = \lim_{H\in\mathcal{E}(I)} \sum_{i\in H} x_i$$

*existiert. $(x_i)_{i\in I}$ ist* **summierbar mit Grenzwert** *$a \in V$, wenn es für alle $\varepsilon > 0$ ein $J = J(\varepsilon) \in \mathcal{E}(I)$ gibt, so daß für alle $H \in \mathcal{E}(I)$ mit $H \supseteq J(\varepsilon)$ gilt:*

$$\|a - s_H\| = \left\| a - \sum_{i\in H} x_i \right\| < \varepsilon.$$

*In diesem Fall schreibt man $a = \sum\limits_{i\in I} x_i = \lim\limits_{H\in\mathcal{E}(I)} \sum\limits_{i\in H} x_i$.*

**Lemma 6.4.8.** *Es sei $(V, \|\cdot\|)$ ein normierter Vektorraum, $I$ und $J$ Mengen und $\varphi : I \to J$ bijektiv. Dann ist eine Familie $(x_j)_{j\in J}$ in $V$ genau dann summierbar, wenn $(x_{\varphi(i)})_{i\in I}$ summierbar ist. In diesem Fall gilt:*

$$\sum_{j\in J} x_j = \sum_{i\in I} x_{\varphi(i)}.$$

*Beweis.*

„$\Rightarrow$“ Es sei $(x_j)_{j\in J}$ summierbar, dann gibt es zu $\varepsilon > 0$ ein $H(\varepsilon) \in \mathcal{E}(J)$, so daß für alle $L \in \mathcal{E}(J)$ mit $H(\varepsilon) \subseteq L$ und für den Grenzwert $x := \sum\limits_{j\in J} x_j$ gilt:

$$\left\| x - \sum_{j\in L} x_j \right\| < \varepsilon.$$

Da $\varphi$ bijektiv ist, sind für jedes solche $L \in \mathcal{E}(J)$ mit $L \supseteq H(\varepsilon)$ auch $\varphi^{-1}(L) \supseteq \varphi^{-1}(H(\varepsilon))$ endliche Teilmengen von $I$ und wegen $\sum\limits_{j\in L} x_j = \sum\limits_{i\in\varphi^{-1}(L)} x_{\varphi(i)}$ ist auch

$$\left\| x - \sum_{j\in L} x_j \right\| = \left\| x - \sum_{i\in\varphi^{-1}(L)} x_{\varphi(i)} \right\| < \varepsilon. \tag{6.4.5}$$

Ferner hat jede endliche Teilmenge $K \in \mathcal{E}(I)$ mit $K \supseteq \varphi^{-1}(H(\varepsilon))$ die Form $K = \varphi^{-1}(L)$ mit $L \in \mathcal{E}(J)$ und $L \supseteq H(\varepsilon)$, daher ist wegen (6.4.5) die Familie $(x_{\varphi(i)})_{i\in I}$ summierbar.

„$\Leftarrow$“ Es sei $(x_{\varphi(i)})_{i\in I}$ summierbar und $\varepsilon > 0$, dann gibt es $L(\varepsilon) \in \mathcal{E}(I)$, so daß für $y := \sum\limits_{i\in I} x_{\varphi(i)}$ und für jedes $K \in \mathcal{E}(I)$ mit $K \supseteq L(\varepsilon)$ gilt $\left\| y - \sum\limits_{i\in K} x_{\varphi(i)} \right\| < \varepsilon$. Jedes $H \in \mathcal{E}(J)$ mit $H \supseteq \varphi(L(\varepsilon))$ hat die Form $H = \varphi(K)$ mit $K \supseteq L(\varepsilon)$, daher gilt:

$$\left\| y - \sum_{j\in H} x_j \right\| = \left\| y - \sum_{i\in K} x_{\varphi(i)} \right\| < \varepsilon$$

also ist $(x_j)_{j\in J}$ summierbar und $y = \sum\limits_{i\in I} x_{\varphi(i)} = \sum\limits_{j\in J} x_j$. □

**Satz 6.4.9.**

*a) Ist $(V, \|\cdot\|)$ ein normierter Vektorraum, $(x_i)_{i\in I}$ eine Familie in $V$, so ist $(\sum\limits_{i\in H} x_i)_{H\in\mathcal{E}(I)}$ genau dann ein Cauchynetz, wenn es für jedes $\varepsilon > 0$ ein $J = J(\varepsilon) \in \mathcal{E}(I)$ gibt, so daß für alle $L \in \mathcal{E}(I)$ mit $L \cap J(\varepsilon) = \emptyset$ gilt:*

$$\left\| \sum_{l\in L} x_l \right\| < \varepsilon \tag{6.4.6}$$

*b) Ist $(X, \|\cdot\|)$ ein Banachraum und $(x_i)_{i\in I}$ eine Familie in $X$, dann ist $(x_i)_{i\in I}$ genau dann summierbar, wenn $\left(\sum_{i\in H} x_i\right)_{H\in\mathcal{E}(I)}$ ein Cauchynetz in $X$ ist.*

*Beweis.*

a)„⇒“ Es sei $\varepsilon > 0$ und $\left(\sum_{i\in H} x_i\right)_{H\in\mathcal{E}(I)}$ ein Cauchynetz in $V$, dann gibt es ein $J = J(\varepsilon) \in \mathcal{E}(I)$, so daß für alle $K, L \in \mathcal{E}(I)$ mit $K \supseteq J(\varepsilon)$ und $L \supseteq J(\varepsilon)$ gilt:

$$\|s_K - s_L\| = \left\|\sum_{i\in K} x_i - \sum_{j\in L} x_j\right\| < \varepsilon \tag{6.4.7}$$

Ist $L \in \mathcal{E}(I)$ mit $L \cap J(\varepsilon) = \emptyset$, so ist

$$s_L = \sum_{i\in L} x_i = s_{L\cup J(\varepsilon)} - s_{J(\varepsilon)} = \sum_{i\in L\cup J(\varepsilon)} x_i - \sum_{i\in J(\varepsilon)} x_i.$$

Wegen $L \cup J(\varepsilon) \supseteq J(\varepsilon)$ folgt nun $\|s_L\| = \|s_{L\cup J(\varepsilon)} - s_{J(\varepsilon)}\| < \varepsilon$ aus (6.4.7).

a)„⇐“ Zu $\varepsilon > 0$ gebe es $J = J(\varepsilon) \in \mathcal{E}(I)$, so daß für alle $L \in \mathcal{E}(I)$ mit $L \cap J(\varepsilon) = \emptyset$ gilt: $\|s_L\| = \left\|\sum_{j\in L} x_j\right\| < \varepsilon$. Sind nun $K, L \in \mathcal{E}(I)$ mit $K \supseteq J(\varepsilon)$ und $L \supseteq J(\varepsilon)$, dann ist $s_L = s_{L\setminus J(\varepsilon)} + s_{J(\varepsilon)}$ und $s_K = s_{K\setminus J(\varepsilon)} + s_{J(\varepsilon)}$ also

$$\|s_K - s_L\| = \|s_{K\setminus J(\varepsilon)} - s_{L\setminus J(\varepsilon)}\| \le \|s_{K\setminus J(\varepsilon)}\| + \|s_{L\setminus J(\varepsilon)}\| < \varepsilon + \varepsilon = 2\varepsilon,$$

also ist $\left(\sum_{i\in H} x_i\right)_{H\in\mathcal{E}(I)}$ ein Cauchynetz in $V$.

b)„⇒“ Es sei $(x_i)_{i\in I}$ summierbar und $x := \lim\limits_{H\in\mathcal{E}(I)} \sum_{i\in H} x_i$ der Grenzwert. Zu $\varepsilon > 0$ gibt es dann $J = J(\varepsilon) \in \mathcal{E}(I)$, so daß für alle $H \in \mathcal{E}(I)$ mit $H \supseteq J(\varepsilon)$ gilt: $\left\|x - \sum_{i\in H} x_i\right\| < \varepsilon$. Daher gilt für alle $K, L \in \mathcal{E}(I)$ mit $K \supseteq J(\varepsilon)$, $L \supseteq J(\varepsilon)$:

$$\|s_K - s_L\| = \left\|\sum_{j\in K} x_j - \sum_{j\in L} x_j\right\| \le \|x - s_K\| + \|x - s_L\| \le 2\varepsilon$$

dh. $\left(\sum_{i\in H} x_i\right)_{H\in\mathcal{E}(I)}$ ist ein Cauchynetz.

b)„⇐“ Nach Satz 6.4.6 ist im Banachraum jedes Cauchynetz konvergent, also auch das Cauchynetz $\left(\sum_{i\in H} x_i\right)_{H\in\mathcal{E}(I)}$, dh. $(x_i)_{i\in I}$ summierbar.

$\square$

**Lemma 6.4.10.** *Es sei $(V, \|\cdot\|)$ ein normierter Vektorraum und $(x_i)_{i\in I}$ eine summierbare Familie in $V$. Dann gilt:*

*a) Für jedes $\varepsilon > 0$ gibt es nur endlich viele Indizes $i_1, ..., i_{N(\varepsilon)} \in I$ mit $\|x_j\| > \varepsilon$ für $j = 1, ..., N(\varepsilon)$.*

*b) $\{\sum\limits_{j\in H} x_j : H \in \mathcal{E}(I)\}$ ist beschränkt.*

*c) $\{i \in I : x_i \neq 0\}$ ist abzählbar.*

*Beweis.* Ist $(x_i)_{i\in I}$ eine summierbare Familie in $V$, so ist $(s_H)_{H\in\mathcal{E}(I)} = \left(\sum\limits_{i\in H} x_i\right)_{H\in\mathcal{E}(I)}$ nach Lemma 6.4.5 ein Cauchynetz in $V$.

a) Nach Satz 6.4.9 gibt es für alle $\varepsilon > 0$ ein $J(\varepsilon) \in \mathcal{E}(I)$, so daß für alle $L \in \mathcal{E}(I)$ mit $L \cap J(\varepsilon) = \emptyset$ gilt: $\|s_L\| = \left\|\sum\limits_{i\in L} x_i\right\| < \varepsilon$. Insbesondere folgt daraus $\|s_{\{i\}}\| = \|x_i\| < \varepsilon$ für alle $i \in I\backslash J(\varepsilon)$ und das zeigt a).

b) Ist $H \in \mathcal{E}(I)$ und $J = J(1) \in \mathcal{E}(I)$ mit $\|s_L\| = \left\|\sum\limits_{j\in L} x_j\right\| < 1$ für alle $L \in \mathcal{E}(I)$ mit $L \cap J = \emptyset$. Dann ist als Supremum einer endlichen Menge

$$C := \sup\left\{\left\|\sum_{j\in K} x_j\right\| : K \subseteq J\right\} < \infty$$

also ist wegen $(H \cap (I\backslash J)) \cap J = \emptyset$ auch

$$\|s_H\| = \left\|\sum_{j\in H} x_j\right\| \leq \left\|\sum_{j\in H\cap J} x_j\right\| + \left\|\sum_{j\in H\cap(I\backslash J)} x_j\right\| \leq C + 1$$

und daher $\sup\{\|s_H\| : H \in \mathcal{E}(I)\} \leq C + 1$.

c) Zu $k \in \mathbb{N}$ sei $H_k := \{i \in I : \|x_i\| \geq \frac{1}{k}\}$. Dann ist $H_k$ nach a) eine endliche Menge und

$$\{i \in I : \|x_i\| \neq 0\} = \bigcup_{k=1}^{\infty} H_k$$

als abzählbare Vereinigung von endlichen Mengen wieder abzählbar.

$\square$

**Satz 6.4.11.** *Es sei $(X, \|\cdot\|)$ ein Banachraum, $(x_i)_{i\in I}$ eine summierbare Familie in $X$ und $J \subseteq I$. Dann ist auch $(x_j)_{j\in J}$ summierbar.*

*Beweis.* Ist $(x_i)_{i\in I}$ summierbar, so ist $(s_H)_{H\in\mathcal{E}(I)} = \left(\sum_{i\in H} x_i\right)_{H\in\mathcal{E}(I)}$ ein Cauchynetz in $X$ und daher gibt es für alle $\varepsilon > 0$ ein $K = K(\varepsilon) \in \mathcal{E}(I)$, so daß $\|s_L\| = \left\|\sum_{j\in L} x_j\right\| < \varepsilon$ für alle $L \in \mathcal{E}(I)$ mit $L \cap K(\varepsilon) = \emptyset$ gilt. Setze nun $\widetilde{K}(\varepsilon) := K(\varepsilon) \cap J \in \mathcal{E}(J)$, dann gilt für jedes $M \in \mathcal{E}(J)$ mit $M \cap \widetilde{K}(\varepsilon) = \emptyset$ auch $M \cap K(\varepsilon) = (M \cap J) \cap K(\varepsilon) = M \cap \widetilde{K}(\varepsilon) = \emptyset$, also ist $\|s_M\| = \left\|\sum_{j\in M} x_j\right\| < \varepsilon$, dh. $(s_M)_{M\in\mathcal{E}(J)} = \left(\sum_{j\in M} x_j\right)_{M\in\mathcal{E}(J)}$ ist ein Cauchynetz in $X$ und da $X$ ein Banachraum ist, ist dieses Cauchynetz auch konvergent und somit $(x_j)_{j\in J}$ eine summierbare Familie. $\square$

**Satz 6.4.12.** *Es sei $(X, \|\cdot\|)$ ein Banachraum, $(x_i)_{i\in I}$ eine summierbare Familie in $X$ und $(I_\lambda)_{\lambda\in\Lambda}$ eine Zerlegung von $I$, dann existiert für jedes $\lambda \in \Lambda$ der Grenzwert $\sum_{i\in I_\lambda} x_i$ und die Familie $\left(\sum_{i\in I_\lambda} x_i\right)_{\lambda\in\Lambda}$ ist summierbar und es gilt:*

$$\sum_{i\in I} x_i = \sum_{\lambda\in\Lambda}\left(\sum_{i\in I_\lambda} x_i\right). \tag{6.4.8}$$

*Beweis.* Ist $(x_i)_{i\in I}$ eine summierbare Familie im Banachraum $X$, so ist nach Satz 6.4.11 für jedes $J \subseteq I$ auch $(x_j)_{j\in J}$ eine summierbare Familie. Insbesondere ist jede Familie $(x_i)_{i\in I_\lambda}$ summierbar und damit existiert für jedes $\lambda \in \Lambda$ der Grenzwert

$$s_\lambda := \sum_{i\in I_\lambda} x_i.$$

Es sei $s := \sum_{i\in I} x_i$ der Grenzwert der summierbaren Familie $(x_i)_{i\in I}$ und $\varepsilon > 0$. Dann gibt es ein $H(\varepsilon) \in \mathcal{E}(I)$, so daß

$$\left\|s - \sum_{i\in J} x_i\right\| < \frac{\varepsilon}{3}$$

für jedes $J \in \mathcal{E}(I)$ mit $J \supseteq H(\varepsilon)$ erfüllt ist. Da $\left(\sum_{i\in H} x_i\right)_{H\in\mathcal{E}(I)}$ ein Cauchynetz ist, können wir nach Satz 6.4.9 sogar (oE. nach Übergang zur gemeinsamen oberen Schranke)

$$\left\|\sum_{i\in L} x_i\right\| < \frac{\varepsilon}{3}$$

für alle $L \in \mathcal{E}(I)$ mit $L \cap H(\varepsilon) = \emptyset$ voraussetzen. Es sei

$$M(\varepsilon) := \{\lambda \in \Lambda : H(\varepsilon) \cap I_\lambda \neq \emptyset\} \in \mathcal{E}(\Lambda),$$

– da $H(\varepsilon)$ endlich ist, ist es auch $M(\varepsilon)$. Ist $K = \{k_1, ..., k_N\} \in \mathcal{E}(\Lambda)$ mit $M(\varepsilon) \subseteq K$, so wähle $J_{k_1} \in \mathcal{E}(I_{k_1})$,...,$J_{k_N} \in \mathcal{E}(I_{k_N})$ mit $H(\varepsilon) \cap I_{k_n} \subseteq J_{k_n}$ und

$$\left\| s_{k_n} - \sum_{j \in J_{k_n}} x_j \right\| < \frac{\varepsilon}{3N}$$

(das geht, da $(x_j)_{j \in I_{k_n}}$ summierbar mit Grenzwert $s_{k_n}$ ist). Nach Definition von $M(\varepsilon)$ ist

$$H(\varepsilon) \subseteq J(\varepsilon) := \bigcup_{\lambda \in M(\varepsilon)} J_\lambda \in \mathcal{E}(I)$$

und

$$L(\varepsilon) := \bigcup_{\lambda \in K \setminus M(\varepsilon)} J_\lambda \in \mathcal{E}(I)$$

erfüllt $L(\varepsilon) \cap H(\varepsilon) = \emptyset$. Die Assoziativität für endliche Summen ergibt

$$\sum_{j \in J(\varepsilon)} x_j = \sum_{\lambda \in M(\varepsilon)} \left( \sum_{j \in J_\lambda} x_j \right)$$

und

$$\sum_{j \in L(\varepsilon)} x_j = \sum_{\lambda \in K \setminus M(\varepsilon)} \left( \sum_{j \in J_\lambda} x_j \right)$$

Somit ist

$$\begin{aligned} \left\| s - \sum_{\lambda \in K} s_\lambda \right\| &= \left\| s - \sum_{\lambda \in K} \left( s_\lambda - \sum_{j \in J_\lambda} x_j + \sum_{j \in J_\lambda} x_j \right) \right\| \\ &= \left\| s - \sum_{\lambda \in M(\varepsilon)} \left( \sum_{j \in J_\lambda} x_j \right) + \sum_{\lambda \in K} \left( s_\lambda - \sum_{j \in J_\lambda} x_j \right) - \sum_{\lambda \in K \setminus M(\varepsilon)} \sum_{j \in J_\lambda} x_j \right\| \\ &\leq \left\| s - \sum_{j \in J(\varepsilon)} x_j \right\| + \sum_{n=1}^{N} \left\| s_{k_n} - \sum_{j \in J_{k_n}} x_j \right\| + \left\| \sum_{j \in L(\varepsilon)} x_j \right\| \leq \frac{\varepsilon}{3} + N \frac{\varepsilon}{3N} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Daher ist $(s_\lambda)_{\lambda \in \Lambda}$ summierbar mit Grenzwert $s = \sum_{i \in I} x_i = \sum_{\lambda \in \Lambda} \left( \sum_{i \in I_\lambda} x_i \right)$. □

**Definition 6.4.13.** *Es sei $(V, \|\cdot\|)$ ein normierter Vektorraum und $(x_i)_{i \in I}$ eine Familie in $V$. Dann heißt $(x_i)_{i \in I}$* **absolut summierbar***, wenn $(\|x_i\|)_{i \in I}$ summierbar in $[0, \infty[$ ist.*

**Lemma 6.4.14.** *Es sei $(V, \|\cdot\|)$ ein normierter Vektorraum und $(x_i)_{i \in I}$ eine Familie in $V$, dann ist $(x_i)_{i \in I}$ genau dann absolut summierbar, wenn*

$$\sup \left\{ \sum_{j \in H} \|x_j\| : H \in \mathcal{E}(I) \right\} < \infty \tag{6.4.9}$$

*ist. In diesem Fall gilt:*

$$\sum_{i\in I} \|x_i\| = \sup\left\{\sum_{j\in H} \|x_j\| : H \in \mathcal{E}(I)\right\}. \tag{6.4.10}$$

*Beweis.*

„$\Rightarrow$“ Ist $(x_i)_{i\in I}$ absolut summierbar, so ist $(\|x_i\|)_{i\in I}$ summierbar, also ist $\left(\sum_{j\in H} \|x_j\|\right)_{H\in\mathcal{E}(I)}$ nach Satz 6.4.9 ein Cauchynetz und daher $\left\{\sum_{j\in H} \|x_j\| : H \in \mathcal{E}(I)\right\} \subseteq \mathbb{R}$ beschränkt, also existiert $\sup\left\{\sum_{j\in H} \|x_j\| : H \in \mathcal{E}(I)\right\} < \infty$.

„$\Leftarrow$“ Ist $s := \sup\left\{\sum_{j\in H} \|x_j\| : H \in \mathcal{E}(I)\right\} < \infty$ und $\varepsilon > 0$, dann gibt es nach Definition des Supremums ein $H(\varepsilon) \in \mathcal{E}(I)$ mit $\sum_{j\in H(\varepsilon)} \|x_j\| \geq s - \varepsilon$. Da $\|x_j\| \geq 0$ ist, gilt

$$\sum_{j\in H(\varepsilon)} \|x_j\| \leq \sum_{j\in L} \|x_j\| \leq s = \sup\left\{\sum_{j\in H} \|x_j\| : H \in \mathcal{E}(I)\right\}$$

für alle $L \in \mathcal{E}(I)$ mit $H(\varepsilon) \subseteq L$ und damit ist $(\|x_i\|)_{i\in I}$ summierbar mit Grenzwert $s = \sup\left\{\sum_{j\in H} \|x_j\| : H \in \mathcal{E}(I)\right\}$. □

**Lemma 6.4.15.** *Es sei $(X, \|\cdot\|)$ ein Banachraum und $(x_i)_{i\in I}$ eine absolut summierbare Familie in $X$. Dann ist $(x_i)_{i\in I}$ summierbar und es gilt:*

$$\left\|\sum_{i\in I} x_i\right\| \leq \sum_{i\in I} \|x_i\|. \tag{6.4.11}$$

*Beweis.* Für $H \in \mathcal{E}(I)$ gilt nach Lemmma 6.4.14:

$$\left\|\sum_{i\in H} x_i\right\| \leq \sum_{i\in H} \|x_i\| \leq \sup\left\{\sum_{i\in L} \|x_i\| : L \in \mathcal{E}(I)\right\} = \sum_{i\in I} \|x_i\|$$

Da $(x_i)_{i\in I}$ absolut summierbar ist, ist $\left(\sum_{i\in H} \|x_i\|\right)_{H\in\mathcal{E}(I)}$ ein Cauchynetz und daher gibt es für jedes $\varepsilon > 0$ nach Satz 6.4.9 ein $H(\varepsilon) \in \mathcal{E}(I)$, so daß $\sum_{i\in L} \|x_i\| < \varepsilon$ für alle $L \in \mathcal{E}(I)$ mit $L \cap H(\varepsilon) = \emptyset$ erfüllt ist. Dann gilt für alle $L \in \mathcal{E}(I)$ mit $L \cap H(\varepsilon) = \emptyset$ auch

$$\left\|\sum_{i\in L} x_i\right\| \leq \sum_{i\in L} \|x_i\| < \varepsilon$$

und damit ist $\left(\sum_{i\in H} x_i\right)_{H\in\mathcal{E}(I)}$ ein Cauchynetz und daher $(x_i)_{i\in I}$ summierbar. Es sei nun $x := \sum_{i\in I} x_i$ der Grenzwert der summierbaren Familie $(x_i)_{i\in I}$, $\varepsilon > 0$ und $K(\varepsilon) \in \mathcal{E}(I)$, so daß für alle $L \in \mathcal{E}(I)$ mit $L \supseteq K(\varepsilon)$ gilt: $\left\|x - \sum_{i\in L} x_i\right\| < \varepsilon$, für solche $L \in \mathcal{E}(I)$ mit $L \supseteq K(\varepsilon)$ gilt

$$\begin{aligned}
\|x\| &= \left\|x - \sum_{i\in L} x_i + \sum_{i\in L} x_i\right\| \le \left\|x - \sum_{i\in L} x_i\right\| + \left\|\sum_{i\in L} x_i\right\| \le \varepsilon + \sum_{i\in L} \|x_i\| \\
&\le \varepsilon + \sup\left\{\sum_{i\in L} \|x_i\| : L \in \mathcal{E}(I), L \supseteq K(\varepsilon)\right\} \le \varepsilon + \sup\left\{\sum_{i\in L} \|x_i\| : L \in \mathcal{E}(I)\right\} \\
&= \varepsilon + \sum_{i\in I} \|x_i\| \qquad (6.4.12)
\end{aligned}$$

und da man (6.4.12) für jedes $\varepsilon > 0$ bekommt, folgt $\|x\| = \left\|\sum_{i\in I} x_i\right\| \le \sum_{i\in I} \|x_i\|$. □

**Satz 6.4.16.** *Es sei $(X, \|\cdot\|)$ ein Banachraum und $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ eine Reihe in $X$, dann sind äquivalent:*

*a)* $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ *ist eine absolut konvergente Reihe.*

*b) Die Familie $(x_n)_{n\in\mathbb{N}}$ ist absolut summierbar.*

*In diesem Fall ist*

$$\sup\left\{\sum_{k=1}^{n} \|x_k\| : n \in \mathbb{N}\right\} = \lim_{n\to\infty} \sum_{k=1}^{n} \|x_k\| \qquad (6.4.13)$$

*der Grenzwert.*

*Beweis.* Wegen $\{1, ..., n\} \in \mathcal{E}(\mathbb{N})$ ist $\sup\left\{\sum_{k=1}^{n} \|x_k\| : n \in \mathbb{N}\right\} \le \sup\left\{\sum_{k\in H} \|x_k\| : H \in \mathcal{E}(\mathbb{N})\right\}$ und weil es zu $H \in \mathcal{E}(\mathbb{N})$ ein $N \in \mathbb{N}$ mit $H \subseteq \{1, ..., N\}$ und damit $\sum_{k\in H} \|x_k\| \le \sum_{k=1}^{N} \|x_k\|$ gibt, folgt $\sup\left\{\sum_{k=1}^{n} \|x_k\| : n \in \mathbb{N}\right\} \ge \sup\left\{\sum_{k\in H} \|x_k\| : H \in \mathcal{E}(\mathbb{N})\right\}$, also insgesamt $\sup\left\{\sum_{k=1}^{n} \|x_k\| : n \in \mathbb{N}\right\} = \sup\left\{\sum_{k\in H} \|x_k\| : H \in \mathcal{E}(\mathbb{N})\right\}$

a)⇒b) Nach Definition ist $\left(\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}}$ genau dann eine absolut konvergente Reihe, wenn $\left(\sum_{k=1}^{n} \|x_k\|\right)_{n\in\mathbb{N}}$ konvergiert. In diesem Fall ist $\left\{\sum_{k=1}^{n} \|x_k\| : n\in\mathbb{N}\right\}$ beschränkt, also $\sup\left\{\sum_{k\in H} \|x_k\| : H\in\mathcal{E}(\mathbb{N})\right\} = \sup\left\{\sum_{k=1}^{n} \|x_k\| : n\in\mathbb{N}\right\} < \infty$. Nach Lemma 6.4.14 ist $(x_k)_{k\in\mathbb{N}}$ dann absolut summierbar.

b)⇒a) Ist umgekehrt $\sup\left\{\sum_{k=1}^{n} \|x_k\| : n\in\mathbb{N}\right\} < \infty$, so ist wegen $\|x_k\| \geq 0$ die Folge $\left(\sum_{k=1}^{n} \|x_k\|\right)_{n\in\mathbb{N}}$ monoton steigend, also konvergent.

□

## 6.5 Konvergenz von Doppelreihen und Cauchyprodukten

**Satz 6.5.1.** *Es sei $(X,\|\cdot\|)$ ein Banachraum und* $\begin{array}{rcl} x:\mathbb{N}_0^2 & \to & X \\ (k,l) & \mapsto & x_{(k,l)} \end{array}$ *eine Abbildung. Dann sind äquivalent:*

*a) Es gibt eine bijektive Abbildung $\varphi:\mathbb{N}_0\to\mathbb{N}_0^2$, so daß die Reihe $\left(\sum_{j=0}^{n} x_{\varphi(j)}\right)_{n\in\mathbb{N}_0}$ absolut konvergent ist.*

*b) $\sup\left\{\sum_{(k,l)\in H} \|x_{(k,l)}\| : H\subseteq\mathbb{N}_0^2 \text{ ist endlich}\right\}$ existiert in $\mathbb{R}$.*

*In diesem Fall existiert der Grenzwert der Doppelreihe*

$$\sum_{(k,l)\in\mathbb{N}_0^2} x_{(k,l)} := \lim_{n\to\infty}\sum_{j=0}^{n} x_{\varphi(j)}$$

*und es gilt*

$$\lim_{n\to\infty}\sum_{j=0}^{n} \|x_{\varphi(j)}\| = \sup\left\{\sum_{(k,l)\in H} \|x_{(k,l)}\| : H\subseteq\mathbb{N}_0^2 \text{ ist endlich}\right\}$$

*Ist nun eine Zerlegung $\mathbb{N}_0^2 = \bigcup_{m\in\mathbb{N}_0} I_m$ gegeben und wählt man für alle abzählbar unendlichen Mengen $I_m$ dieser Zerlegung eine bijektive Abbildung $\varphi_m:\mathbb{N}_0\to I_m$, so sind alle Reihen*

$\left(\sum_{j=0}^{n} x_{\varphi_m(j)}\right)_{n\in\mathbb{N}_0}$ *absolut konvergent. Setzt man nun*

$$\sum_{(k,l)\in I_m} x_{(k.l)} := \lim_{n\to\infty} \sum_{j=0}^{n} x_{\varphi_m(j)}$$

*für die abzählbar unendlichen Mengen $I_m$ in dieser Zerlegung und $\sum\limits_{(k,l)\in I_m} x_{(k.l)}$ für die Summe im Fall einer endlichen Menge $I_m$, dann ist auch die Reihe*

$$\left(\sum_{m=0}^{M}\left(\sum_{(k,l)\in I_m} x_{(k.l)}\right)\right)_{M\in\mathbb{N}_0}$$

*absolut konvergent und es gilt*

$$\sum_{(k,l)\in\mathbb{N}_0^2} x_{(k,l)} := \lim_{n\to\infty}\sum_{j=0}^{n} x_{\varphi(j)} = \lim_{M\to\infty}\left(\sum_{m=0}^{M}\left(\sum_{(k,l)\in I_m} x_{(k.l)}\right)\right).$$

*Insbesondere ist dann*

- *für jedes $k\in\mathbb{N}_0$ die Reihe $\left(\sum_{l=0}^{n} x_{(k,l)}\right)_{n\in\mathbb{N}_0}$ absolut konvergent*
- *für jedes $l\in\mathbb{N}_0$ die Reihe $\left(\sum_{k=0}^{m} x_{(k,l)}\right)_{m\in\mathbb{N}_0}$ absolut konvergent*

*und die Reihen der Grenzwerte $\left(\sum_{k=0}^{m}\left(\sum_{l=0}^{\infty} x_{(k,l)}\right)\right)_{m\in\mathbb{N}_0}$ und $\left(\sum_{l=0}^{n}\left(\sum_{k=0}^{\infty} x_{(k,l)}\right)\right)_{n\in\mathbb{N}_0}$ sind absolut konvergent und es gilt:*

$$\sum_{k=0}^{\infty}\left(\sum_{l=0}^{\infty} x_{(k,l)}\right) = \sum_{l=0}^{\infty}\left(\sum_{k=0}^{\infty} x_{(k,l)}\right). \tag{6.5.1}$$

*Beweis.*

a)⇒b) Gibt es eine bijektive Abbildung $\varphi:\mathbb{N}_0\to\mathbb{N}_0^2$, so daß die Reihe $\left(\sum_{j=0}^{n} x_{\varphi(j)}\right)_{n\in\mathbb{N}_0}$ absolut konvergent ist, dann ist $(\|x_{\varphi(j)}\|)_{j\in\mathbb{N}}$ nach Satz 6.4.16 (absolut) summierbar. Da auch $\varphi^{-1}:\mathbb{N}_0^2\to\mathbb{N}_0$ bijektiv ist, ist $(x_{(k,l)})_{(k,l)\in\mathbb{N}_0^2}$ nach Lemma 6.4.8 (absolut) summierbar, woraus

$$\sup\left\{\sum_{(k,l)\in H}\|x_{(k,l)}\| : H\subseteq\mathbb{N}_0^2 \text{ ist endlich}\right\} \text{ existiert in } \mathbb{R}$$

mit Lemma 6.4.14 folgt.

b)⇒a) Ist $\varphi : \mathbb{N}_0 \to \mathbb{N}_0^2$ bijektiv und

$$\sup \left\{ \sum_{(k,l)\in H} \|x_{(k,l)}\| : H \subseteq \mathbb{N}_0^2 \text{ ist endlich} \right\} \text{existiert in } \mathbb{R}$$

so ist die Familie $(\|x_{(k,l)}\|)_{(k,l)\in\mathbb{N}_0^2}$ (absolut) summierbar, also ist auch $(\|x_{\varphi(i)}\|)_{i\in\mathbb{N}_0}$ nach Lemma 6.4.8 (absolut) summierbar, daher konvergiert die Reihe $\left(\sum_{j=0}^{n} x_{\varphi(j)}\right)_{n\in\mathbb{N}_0}$ nach Satz 6.4.16 absolut.

Jede der Bedingungen stellt also sicher, daß die Familie $(x_{(k,l)})_{(k,l)\in\mathbb{N}_0^2}$ absolut summierbar ist. Nach Lemma 6.4.15 ist dann auch $(x_{(k,l)})_{(k,l)\in\mathbb{N}_0^2}$ summierbar; nach Satz 6.4.11 sind alle Familien $(x_{(k,l)})_{(k,l)\in I_m}$ mit $I_m \subseteq \mathbb{N}_0^2$ summierbar. Ist nun $\mathbb{N}_0^2 = \bigcup_{m\in\mathbb{N}_0} I_m$ eine Zerlegung von $\mathbb{N}_0^2$, so zeigt Satz 6.4.12 die fehlenden Gleichheiten in Satz 6.5.1. □

**Definition 6.5.2.** *Ein Banachraum $(X, \|\cdot\|)$, der gleichzeitig ein Ring $(X, +, \cdot)$ ist, heißt* **Banachalgebra***, wenn $\|x \cdot y\| \le \|x\|\ \|y\|$ für alle $x, y \in X$ gilt. Sind $\left(\sum_{k=0}^{n} x_k\right)_{n\in\mathbb{N}_0}$ und $\left(\sum_{k=0}^{n} y_k\right)_{n\in\mathbb{N}_0}$ Reihen in einer Banachalgebra $(X, \|\cdot\|)$, dann heißt die Reihe $\left(\sum_{k=0}^{n} c_k\right)_{n\in\mathbb{N}_0}$ mit dem Reihenglied*

$$c_k := \sum_{l=0}^{k} x_l y_{k-l} \tag{6.5.2}$$

*das* **Cauchyprodukt** *von $\left(\sum_{k=0}^{n} x_k\right)_{n\in\mathbb{N}_0}$ und $\left(\sum_{k=0}^{n} y_k\right)_{n\in\mathbb{N}_0}$.*

**Bemerkung 6.5.3.** $(\mathbb{R}, |\cdot|)$ ist eine $\mathbb{R}-$Banachalgebra und $(\mathbb{C}, |\cdot|)$ ist eine $\mathbb{C}-$Banachalgebra. Ebenso wie in Lemma 5.2.1 konvergiert mit zwei Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ in einer Banachalgebra $X$ auch die Produktfolge $(x_n \cdot y_n)_{n\in\mathbb{N}}$ und es gilt

$$\lim_{n\to\infty} (x_n \cdot y_n) = \left(\lim_{n\to\infty} x_n\right) \cdot \left(\lim_{n\to\infty} y_n\right).$$

**Satz 6.5.4.** *Sind $\left(\sum_{k=0}^{n} x_k\right)_{n\in\mathbb{N}_0}$ und $\left(\sum_{k=0}^{n} y_k\right)_{n\in\mathbb{N}_0}$ absolut konvergente Reihen in einer Banachalgebra $(X, \|\cdot\|)$, dann ist auch das Cauchyprodukt $\left(\sum_{k=0}^{n} c_k\right)_{n\in\mathbb{N}_0}$ von $\left(\sum_{k=0}^{n} x_k\right)_{n\in\mathbb{N}_0}$ und $\left(\sum_{k=0}^{n} y_k\right)_{n\in\mathbb{N}_0}$ absolut konvergent und es gilt:*

$$\sum_{k=0}^{\infty} c_k = \left(\sum_{k=0}^{\infty} x_k\right) \cdot \left(\sum_{k=0}^{\infty} y_k\right). \tag{6.5.3}$$

*Beweis.* Für jede endliche Menge $H \subseteq \mathbb{N}_0^2$ gibt es endliche Mengen $K, L \subseteq \mathbb{N}_0$ mit $H \subseteq K \times L$ und damit ist

$$\begin{aligned}\sum_{(k,l)\in H} \|x_k \cdot y_l\| &\leq \sum_{(k,l)\in H} \|x_k\| \cdot \|y_l\| \leq \sum_{(k,l)\in K\times L} \|x_k\| \cdot \|y_l\| = \left(\sum_{k\in K} \|x_k\|\right)\left(\sum_{l\in L} \|y_l\|\right) \\ &\leq \left(\sum_{k=0}^{\infty} \|x_k\|\right)\left(\sum_{l=0}^{\infty} \|y_l\|\right)\end{aligned}$$

und damit Satz 6.5.1 anwendbar. Mit der Zerlegung $\mathbb{N}_0^2 = \bigcup_{m\in\mathbb{N}_0} I_m$ mit

$$I_m := \left\{(k,l) \in \mathbb{N}_0^2 : k + l = m\right\} \tag{6.5.4}$$

gibt Satz 6.5.1 die Konvergenz des Cauchyprodukts. □

**Lemma 6.5.5.** *Es sei $X$ eine $\mathbb{C}-$Banachalgebra, $(b_n)_{n\in\mathbb{N}_0}$ und $(c_n)_{n\in\mathbb{N}_0}$ Folgen in $X$, $a \in \mathbb{C}$ und $r > 0$, so daß für die Funktionen*

$$\begin{aligned} f : K(a,r) &\to X & \quad \textit{und} \quad g : K(a,r) &\to X \\ z &\mapsto \sum_{n=0}^{\infty} b_n (z-a)^n & z &\mapsto \sum_{n=0}^{\infty} c_n (z-a)^n \end{aligned}$$

*die Potenzreihen für $|z-a| < r$ konvergieren. Dann konvergiert die Potenzreihe*

$$\left(\sum_{n=0}^{N} \left(\sum_{k=0}^{n} b_k c_{n-k}\right)(z-a)^n\right)_{N\in\mathbb{N}_0}$$

*zum Cauchyprodukt für alle $|z-a| < r$ und es gilt*

$$f(z)g(z) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} b_k c_{n-k}\right)(z-a)^n \tag{6.5.5}$$

*für $|z-a| < r$.*

*Beweis.* Es sei $s \in ]0, r[$, dann gibt es nach Lemma 6.3.5 ein $C \in [0, \infty[$ mit $\|b_n\| \leq \dfrac{C}{s^n}$ und $\|c_n\| \leq \dfrac{C}{s^n}$ für alle $n \in \mathbb{N}_0$. Somit ist

$$\left\|\sum_{k=0}^{n} b_k c_{n-k}\right\| \leq \sum_{k=0}^{n} \|b_k\| \cdot \|c_{n-k}\| \leq (n+1)\frac{C^2}{s^n}$$

und daher

$$\left\|\sum_{k=0}^{n} b_k c_{n-k}\right\|^{\frac{1}{n}} \leq \frac{\sqrt[n]{(n+1)C^2}}{s} \stackrel{n\to\infty}{\longrightarrow} \frac{1}{s}. \tag{6.5.6}$$

Da (6.5.6) für alle $s \in ]0, r[$ erfüllt wird, folgt

$$\overline{\lim_{n\to\infty}} \left\| \sum_{k=0}^{n} b_k c_{n-k} \right\|^{\frac{1}{n}} \leq \frac{1}{r} \tag{6.5.7}$$

also ist der Konvergenzradius der Potenzreihe $\left( \sum_{n=0}^{N} \left( \sum_{k=0}^{n} b_k c_{n-k} \right) (z-a)^n \right)_{N\in\mathbb{N}_0}$ mindestens $r$. Auf $K(a,r)$ sind also die Funktionen $f$ und $g$ wohldefiniert und wegen

$$f(z) = \lim_{N\to\infty} \sum_{n=0}^{N} b_n (z-a)^n \quad \text{und} \quad g(z) = \lim_{N\to\infty} \sum_{n=0}^{N} c_n (z-a)^n$$

ist nach Bemerkung 6.5.3

$$f(z)g(z) = \lim_{N\to\infty} \left( \left( \sum_{n=0}^{N} b_n (z-a)^n \right) \left( \sum_{n=0}^{N} c_n (z-a)^n \right) \right)$$

für $|z-a| < r$. Da die Potenzreihe $\left( \sum_{n=0}^{N} \left( \sum_{k=0}^{n} b_k c_{n-k} \right) (z-a)^n \right)_{N\in\mathbb{N}_0}$ einen Konvergenzradius $\geq r$ hat, konvergiert für $N < M$

$$\begin{aligned}
&\left\| \left( \sum_{n=0}^{N} b_n (z-a)^n \right) \left( \sum_{n=0}^{N} c_n (z-a)^n \right) - \sum_{n=0}^{M} \left( \sum_{k=0}^{n} b_k c_{n-k} \right) (z-a)^n \right\| \\
&= \left\| \sum_{n=0}^{2N} \left( \sum_{k=0}^{\min\{n,N\}} b_k c_{n-k} \right) (z-a)^n - \sum_{n=0}^{M} \left( \sum_{k=0}^{n} b_k c_{n-k} \right) (z-a)^n \right\| \\
&\leq \sum_{n=N+1}^{M} \left( \sum_{k=0}^{n} \|b_k\| \cdot \|c_{n-k}\| \right) |z-a|^n \overset{N,M\to\infty}{\longrightarrow} 0
\end{aligned}$$

für $|z-a| < r$ (als Reihenrest der absolut konvergenten Potenzreihe) und daher folgt

$$f(z)g(z) = \sum_{n=0}^{\infty} \left( \sum_{k=0}^{n} b_k c_{n-k} \right) (z-a)^n$$

für alle $|z-a| < r$. $\square$

## 6.6 Exponentialfunktion, Sinus und Cosinus

**Lemma 6.6.1.** *Für jedes $z \in \mathbb{C}$ ist die Reihe* $\left(\sum_{k=0}^{n} \frac{z^k}{k!}\right)_{n\in\mathbb{N}_0}$ *absolut konvergent.*

*Beweis.* Für jedes $z \in \mathbb{C}\backslash\{0\}$ ist

$$\left|\frac{\frac{z^{k+1}}{(k+1)!}}{\frac{z^k}{k!}}\right| = \frac{|z|}{k+1} \xrightarrow{k\to\infty} 0$$

und daher ist $\left(\sum_{k=0}^{n} \frac{z^k}{k!}\right)_{n\in\mathbb{N}_0}$ nach dem Quotientenkriterium absolut konvergent. □

**Definition 6.6.2.**

- *Die Funktion* $\begin{array}{rcl} \exp : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & e^z := \sum_{k=0}^{\infty} \frac{z^k}{k!} \end{array}$ *heißt (komplexe)* **Exponentialfunktion**.

- $\begin{array}{rcl} \cos : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & \frac{e^{iz}+e^{-iz}}{2} \end{array}$ *heißt (komplexe)* **Cosinusfunktion**.

- $\begin{array}{rcl} \sin : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & \frac{e^{iz}-e^{-iz}}{2i} \end{array}$ *heißt (komplexe)* **Sinusfunktion**.

- $e := e^1 = \sum_{k=0}^{\infty} \frac{1}{k!}$ *heißt* **Eulersche Zahl**.

**Satz 6.6.3.**

- $e^0 = 1$.

- *Für alle $w, z \in \mathbb{C}$ gilt die* **Funktionalgleichung für die Exponentialfunktion***:*

$$e^{w+z} = e^w \cdot e^z. \tag{6.6.1}$$

- *Für alle $n \in \mathbb{N}$ gilt:* $e^n = \underbrace{e \cdots e}_{n-mal}$.

- *Für alle $z \in \mathbb{C}$ ist*

$$\begin{aligned} e^{-z} &= (e^z)^{-1} && (6.6.2)\\ e^z &\neq 0 && (6.6.3)\\ e^{\overline{z}} &= \overline{e^z} && (6.6.4)\\ |e^z| &= e^{\operatorname{Re} z} && (6.6.5)\\ e^z &= \lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n = \lim_{n\to\infty}\frac{1}{\left(1-\frac{z}{n}\right)^n} && (6.6.6) \end{aligned}$$

- *Für alle $x \in \mathbb{R}$ ist $e^x \in \mathbb{R}$ und*

$$\begin{aligned} |e^{ix}| &= 1 && (6.6.7)\\ e^x &> 0 && (6.6.8) \end{aligned}$$

*Beweis.*

- Einsetzen von $z = 0$ ergibt $e^0 = 1 + \sum_{k=1}^{\infty} \frac{0^k}{k!} = 1$.

- Nach Lemma 6.6.1 ist die Exponentialreihe auf ganz $\mathbb{C}$ absolut konvergent, daher sind die beiden Reihen $\left(\sum_{k=0}^{n} \frac{w^k}{k!}\right)_{n\in\mathbb{N}_0}$ und $\left(\sum_{k=0}^{n} \frac{z^k}{k!}\right)_{n\in\mathbb{N}_0}$ absolut konvergent, deshalb ist nach Satz 6.5.4 auch das Cauchyprodukt $\left(\sum_{k=0}^{n} c_k\right)_{n\in\mathbb{N}_0}$ dieser Reihen absolut konvergent. In diesem Fall ist das $k$-te Reihenglied $c_k$ des Cauchyprodukts gerade

$$c_k = \sum_{l=0}^{k} \frac{z^{k-l}}{(k-l)!}\frac{w^l}{l!} = \frac{1}{k!}\sum_{l=0}^{k}\binom{k}{l} w^l z^{k-l} = \frac{1}{k!}(w+z)^k$$

Nach Satz 6.5.4 und der Definition der Exponentialfunktion gilt:

$$e^{w+z} = \sum_{k=0}^{\infty}\frac{(w+z)^k}{k!} = \left(\sum_{k=0}^{\infty}\frac{w^k}{k!}\right)\cdot\left(\sum_{k=0}^{\infty}\frac{z^k}{k!}\right) = e^w \cdot e^z.$$

- Mit Induktion folgt nun $e^n = \underbrace{e\cdots e}_{n-mal}$. und aus $1 = e^0 = e^{z-z} = e^z e^{-z}$ folgt (6.6.2). Ferner ist dann $e^z \neq 0$ für jedes $z \in \mathbb{C}$, denn sonst ist $1 = e^z e^{-z} = 0$.

- Für die Partialsummen von $e^{\overline{z}}$ und $\overline{e^z}$ gilt:

$$\sum_{k=0}^{n}\frac{(\overline{z})^k}{k!} = \overline{\sum_{k=0}^{n}\frac{z^k}{k!}}$$

und da mit jeder konvergenten Folge $(w_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ mit $w = \lim_{n\to\infty} w_n$ nach Lemma 5.4.8 auch $(\overline{w_n})_{n\in\mathbb{N}}$ konvergiert mit $\overline{w} = \lim_{n\to\infty}\overline{w_n}$, folgt daraus (6.6.4).

- Für alle $x \in \mathbb{R}$ konvergiert $\left(\sum_{k=0}^{n}\frac{x^k}{k!}\right)_{n\in\mathbb{N}}$ und besitzt einen reellen Grenzwert. Für $x \geq 0$ ist $e^x = 1 + \sum_{k=1}^{\infty}\frac{x^k}{k!} \geq 1 > 0$ und für $x < 0$ ist $1 = e^0 = e^{-x}e^x$ mit $e^{-x} \geq 1$, also $e^x > 0$.

- Nach Definition des Absolutbetrags gilt:

$$|e^z| = \sqrt{e^z \overline{e^z}} = \sqrt{e^z e^{\overline{z}}} = \sqrt{e^{z+\overline{z}}} = \sqrt{e^{2\operatorname{Re} z}} = e^{\operatorname{Re} z}.$$

Insbesondere ist $|e^{ix}| = e^0 = 1$ für alle $x \in \mathbb{R}$.

- Da die Exponentialreihe absolut konvergent ist, können wir zu $\varepsilon > 0$ ein $K \in \mathbb{N}$ wählen mit $\displaystyle\sum_{k=K}^{\infty} \frac{|z|^k}{k!} < \frac{\varepsilon}{3}$. Wegen

$$\begin{aligned} \binom{n}{k} \frac{1}{n^k} &= \frac{n!}{k!(n-k)!n^k} = \frac{n(n-1)\cdots(n-k+1)}{k!n^k} \\ &= \frac{1}{k!} 1 \cdot (1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n}) \end{aligned}$$

gilt $\displaystyle\binom{n}{k} \frac{1}{n^k} \le \frac{1}{k!}$ für alle $k \in \mathbb{N}$. Bei gegebenem $k \in \mathbb{N}$ ist $\displaystyle\lim_{n\to\infty} \binom{n}{k} \frac{1}{n^k} = \frac{1}{k!}$; insbesondere gibt es $N \in \mathbb{N}$, so daß für alle $n \ge N$ gilt:

$$\sum_{k=0}^{K-1} \left| \binom{n}{k} \frac{z^k}{n^k} - \frac{z^k}{k!} \right| < \frac{\varepsilon}{3}.$$

Für alle $n \ge N$ gilt dann

$$\begin{aligned} &\left| \left(1 + \frac{z}{n}\right)^n - \sum_{k=0}^{\infty} \frac{z^k}{k!} \right| = \left| \sum_{k=0}^{n} \binom{n}{k} \frac{z^k}{n^k} - \sum_{k=0}^{\infty} \frac{z^k}{k!} \right| \\ &\le \sum_{k=0}^{K-1} \left| \binom{n}{k} \frac{z^k}{n^k} - \frac{z^k}{k!} \right| + \sum_{k=K}^{n} \binom{n}{k} \frac{|z|^k}{n^k} + \sum_{k=K}^{\infty} \frac{|z|^k}{k!} \le \frac{\varepsilon}{3} + \sum_{k=K}^{n} \frac{|z|^k}{k!} + \sum_{k=K}^{\infty} \frac{|z|^k}{k!} \le \varepsilon, \end{aligned}$$

was $e^z = \displaystyle\lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n$ zeigt. Wendet man diese Formel für $-z$ statt $z$ an, so folgt $e^{-z} = \displaystyle\lim_{n\to\infty} \left(1 - \frac{z}{n}\right)^n$, woraus mit (6.6.3) und den Rechenregeln für Grenzwerte $e^z = (e^{-z})^{-1} = \dfrac{1}{\displaystyle\lim_{n\to\infty} \left(1 - \frac{z}{n}\right)^n}$ folgt. □

**Satz 6.6.4.**

- *Für alle $z \in \mathbb{C}$ gilt:*

$$\cos z + i \sin z = e^{iz} \qquad \textbf{(Eulersche Formel)} \tag{6.6.9}$$

$$(\cos z)^2 + (\sin z)^2 = 1 \tag{6.6.10}$$

$$\sin z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} \tag{6.6.11}$$

$$\cos z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!} \tag{6.6.12}$$

$$\sin(-z) = -\sin z \tag{6.6.13}$$

$$\cos(-z) = \cos z \tag{6.6.14}$$

$$\cos(2z) = (\cos z)^2 - (\sin z)^2 \tag{6.6.15}$$

$$\sin(2z) = 2(\sin z)(\cos z) \tag{6.6.16}$$

- *Für alle $w, z \in \mathbb{C}$ gilt:*

$$\cos(w+z) = (\cos w)(\cos z) - (\sin w)(\sin z) \tag{6.6.17}$$

$$\sin(w+z) = (\sin w)(\cos z) + (\cos w)(\sin z) \tag{6.6.18}$$

$$\cos w - \cos z = 2 \sin\left(\frac{z+w}{2}\right) \sin\left(\frac{z-w}{2}\right) \tag{6.6.19}$$

*Beweis.* Nach Definition von Sinus und Cosinus gilt:

$$\begin{aligned}
\cos z + i \sin z &= \frac{e^{iz} + e^{-iz}}{2} + i\frac{e^{iz} - e^{-iz}}{2i} = e^{iz} \\
(\cos z)^2 + (\sin z)^2 &= \left(\frac{e^{iz} + e^{-iz}}{2}\right)^2 + \left(\frac{e^{iz} - e^{-iz}}{2i}\right)^2 \\
&= \frac{e^{2iz} + 2e^{iz}e^{-iz} + e^{-2iz}}{4} - \frac{e^{2iz} - 2e^{iz}e^{-iz} + e^{-2iz}}{4} = 1
\end{aligned}$$

Da die Exponentialreihe absolut konvergiert, können wir nach Satz 6.2.11 die Reihen für $e^{iz}$ und $e^{-iz}$ folgendermaßen umordnen:

$$\begin{aligned}
\sin z &= \frac{e^{iz} - e^{-iz}}{2i} = \frac{1}{2i}\sum_{k=0}^{\infty} \frac{(iz)^k}{k!} - \frac{1}{2i}\sum_{k=0}^{\infty} \frac{(-iz)^k}{k!} \\
&= \frac{1}{2i}\sum_{k=0}^{\infty} \frac{(iz)^{4k}}{(4k)!} + \frac{1}{2i}\sum_{k=0}^{\infty} \frac{(iz)^{4k+1}}{(4k+1)!} + \frac{1}{2i}\sum_{k=0}^{\infty} \frac{(iz)^{4k+2}}{(4k+2)!} + \frac{1}{2i}\sum_{k=0}^{\infty} \frac{(iz)^{4k+3}}{(4k+3)!} \\
&\quad -\frac{1}{2i}\sum_{k=0}^{\infty} \frac{(-iz)^{4k}}{(4k)!} - \frac{1}{2i}\sum_{k=0}^{\infty} \frac{(-iz)^{4k+1}}{(4k+1)!} - \frac{1}{2i}\sum_{k=0}^{\infty} \frac{(-iz)^{4k+2}}{(4k+2)!} - \frac{1}{2i}\sum_{k=0}^{\infty} \frac{(-iz)^{4k+3}}{(4k+3)!}
\end{aligned}$$

Wegen $i^{4k} = 1$, $i^{4k+1} = i$, $i^{4k+2} = -1$ und $i^{4k+3} = -i$ heben sich in der letzten Gleichung gerade alle Reihengrenzwerte zu den geraden Potenzen von $z$ weg und es bleibt noch

$$\sin z = \sum_{k=0}^{\infty} \frac{z^{4k+1}}{(4k+1)!} - \sum_{k=0}^{\infty} \frac{z^{4k+3}}{(4k+3)!} = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!},$$

letzteres wieder nach Umordnen der beiden absolut konvergenten Reihen $\sum_{k=0}^{\infty} \frac{z^{4k+1}}{(4k+1)!}$ und $\sum_{k=0}^{\infty} \frac{z^{4k+3}}{(4k+3)!}$. Analog ergibt sich

$$\cos z = \frac{e^{iz} + e^{-iz}}{2} = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}.$$

Setzt man $-z$ statt $z$ in die Potenzreihendarstellung (6.6.11) des Sinus bzw. (6.6.12) des Cosinus ein, so ergeben sich (6.6.13) bzw. (6.6.14). Für $w, z \in \mathbb{C}$ gilt nach (6.6.1) und (6.6.9):

$$\begin{aligned} e^{i(w+z)} &= \cos(w+z) + i\sin(w+z) = e^{iw}e^{iz} = (\cos z + i\sin z)(\cos w + i\sin w) \\ &= (\cos w\cos z - \sin w\sin z) + i(\cos w\sin z + \sin w\cos z) \end{aligned} \tag{6.6.20}$$

und durch Vergleich von Real- und Imaginärteil folgen (6.6.17) und (6.6.18) für $z, w \in \mathbb{R}$. Im allgemeinen Fall betrachte noch

$$\begin{aligned} e^{-i(w+z)} &= \cos(w+z) - i\sin(w+z) = e^{-iw}e^{-iz} = (\cos z - i\sin z)(\cos w - i\sin w) \\ &= (\cos w\cos z - \sin w\sin z) - i(\cos w\sin z + \sin w\cos z) \end{aligned} \tag{6.6.21}$$

und dann folgen (6.6.17) und (6.6.18) durch Addition bzw. Subtraktion von (6.6.20) und (6.6.21). (6.6.15), (6.6.16) und sind Spezialfälle von (6.6.17) und (6.6.18) mit $w = z$. Mit $z = \frac{z+w}{2} + \frac{z-w}{2}$ und $w = \frac{z+w}{2} - \frac{z-w}{2}$ ergibt (6.6.17) unter Verwendung von (6.6.13):

$$\begin{aligned} \cos(w) &= \cos\left(\frac{z+w}{2}\right)\cos\left(\frac{z-w}{2}\right) + \sin\left(\frac{z+w}{2}\right)\sin\left(\frac{z-w}{2}\right) \\ \cos(z) &= \cos\left(\frac{z+w}{2}\right)\cos\left(\frac{z-w}{2}\right) - \sin\left(\frac{z+w}{2}\right)\sin\left(\frac{z-w}{2}\right) \end{aligned}$$

und (6.6.19) folgt. □

# Kapitel 7

# Lineare Unabhängigkeit, Basen und Dimension

## 7.1 Lineare Unabhängigkeit, Erzeugendensysteme und Basen

**Definition 7.1.1.** *Es sei $V$ ein $K$-Vektorraum.*

- $v_1, ..., v_n \in V$ *heißen* **linear unabhängig**, *wenn für alle* $(\lambda_1, ..., \lambda_n) \in K^n$ *die Folgerung*

$$\lambda_1 v_1 + ... + \lambda_n v_n = \mathbf{0} \Rightarrow \lambda_1 = ... = \lambda_n = 0 \tag{7.1.1}$$

  *gilt.*

- $W \subseteq V$ *heißt linear unabhängig, wenn jede Wahl von endlich vielen paarweise verschiedenen Vektoren aus $W$ linear unabhängig ist.*

- $W \subseteq V$ *heißt* **linear abhängig**, *wenn $W$ nicht linear unabhängig ist.*

- $W \subseteq V$ *heißt ein* **Erzeugendensystem** *von $V$, wenn* $\mathrm{lin}(W) = V$ *ist.*

- $\mathcal{B} \subseteq V$ *heißt eine* **Basis** *von $V$, wenn $\mathcal{B}$ ein linear unabhängiges Erzeugendensystem ist.*

- $V$ *heißt ein* **endlich dimensionaler** *$K$-Vektorraum (in Zeichen* $\dim_K V < \infty$*), wenn es eine Basis $\mathcal{B}$ von $V$ gibt, die nur endlich viele Elemente besitzt.*

**Beispiel 7.1.2.**

a) $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ sind in $\mathbb{R}^2$ linear unabhängig, denn für $\lambda_1, \lambda_2 \in \mathbb{R}$ folgt aus $\lambda_1 v_1 + \lambda_2 v_2 = \begin{pmatrix} \lambda_1 + \lambda_2 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ zuerst $\lambda_2 = 0$ aus der unteren Gleichung und dann $\lambda_1 = 0$ nach Einsetzen in die obere Gleichung.

b) „Dieselben“ Vektoren $v_1 = 1$ und $v_2 = 1 + i$ in $\mathbb{C}$ sind in $\mathbb{C}$ linear abhängig, denn $(-1 - i) \cdot 1 + 1 \cdot (1 + i) = 0$.

c) Die Frage, wann ein einzelner Vektor linear abhängig oder linear unabhängig ist, läßt sich ganz einfach entscheiden: In jedem $K$-Vektorraum $V$ ist der Nullvektor $\mathbf{0}$ linear abhängig und jeder Vektor $v \neq \mathbf{0}$ linear unabhängig.

**Satz 7.1.3.** *Ist $\{\mathbf{0}\} \neq V$ ein $K$-Vektorraum, so gibt es (mindestens) eine Basis $\mathcal{B}$ von $V$.*

*Beweis.* Da $V \neq \{\mathbf{0}\}$ ist, gibt es $v \in V \setminus \{\mathbf{0}\}$ und damit ist $\{v\}$ eine linear unabhängige Teilmenge von $V$, also ist

$$\mathcal{M} := \{W \subseteq V : W \text{ linear unabhängig } \} \neq \emptyset.$$

- Wir zeigen nun, daß $(\mathcal{M}, \subseteq)$ induktiv geordnet ist. Es sei dazu $(W_i)_{i \in I}$ eine total geordnete Teilmenge von $\mathcal{M}$, dann ist $W := \bigcup_{i \in I} W_i$ eine obere Schranke in $\mathcal{M}$: Sind nämlich $w_1, ..., w_m \in W$ paarweise verschieden, so gibt es $i_1, ..., i_m \in I$ mit $w_1 \in W_{i_1}$,...,$w_m \in W_{i_m}$. Da die Familie $(W_i)_{i \in I}$ total geordnet bzgl. $\subseteq$ ist, dürfen wir oE. $W_{i_1}, ..., W_{i_{m-1}} \subseteq W_{i_m}$ annehmen. Dann sind $w_1, ..., w_m \in W_{i_m}$, also linear unabhängig.
- Da wir nun $(\mathcal{M}, \subseteq)$ als induktiv geordnete Menge nachgewiesen haben, gibt es nach dem Zornschen Lemma 1.4.13 (mindestens) ein maximales Element $\mathcal{B}$ in $(\mathcal{M}, \subseteq)$.
- Angenommen $\operatorname{lin}(\mathcal{B}) \neq V$, dann finden wir ein $v \in V \setminus \operatorname{lin}(\mathcal{B})$ und dann sind $v + \lambda_1 w_1 + ... + \lambda_m w_m \neq \mathbf{0}$ für alle $m \in \mathbb{N}$, $w_1, ..., w_m \in \mathcal{B}$ und $(\lambda_1, ..., \lambda_m) \in K^m$. Dies zeigt aber, daß auch $\mathcal{B} \cup \{v\}$ linear unabhängig ist, was im Widerspruch zur Maximalität von $\mathcal{B}$ steht.

Jedes maximale Element $\mathcal{B}$ von $(\mathcal{M}, \subseteq)$ ist also Basis von $V$. □

**Satz 7.1.4.** *Es sei $V$ ein $K$-Vektorraum und $\mathcal{B} \subseteq V$, dann sind äquivalent:*

*a) $\mathcal{B}$ ist eine Basis von $V$.*

*b) $V = \operatorname{lin}(\mathcal{B})$ und für alle $w \in \mathcal{B}$ ist $V \neq \operatorname{lin}(\mathcal{B} \setminus \{w\})$.*

*c) Zu jedem $v \in V \setminus \{\mathbf{0}\}$ gibt es eindeutiges $n \in \mathbb{N}$, $v_1, ..., v_n \in \mathcal{B}$ paarweise verschieden und $\lambda_1, ..., \lambda_n \in K \setminus \{0\}$ mit $v = \lambda_1 v_1 + ... + \lambda_n v_n$.*

*d) $\mathcal{B}$ ist linear unabhängig und für alle $v \in V \setminus \mathcal{B}$ ist $\mathcal{B} \cup \{v\}$ linear abhängig.*

*Beweis.*

a)⇒ b) (durch ¬ b) ⇒ ¬ a)): Ist $V = \operatorname{lin}(\mathcal{B})$ und $w \in \mathcal{B}$ mit $V = \operatorname{lin}(\mathcal{B} \setminus \{w\})$, dann gibt es $n \in \mathbb{N}$, $v_1, ..., v_n \in \mathcal{B} \setminus \{w\}$ paarweise verschieden und $\lambda_1, ..., \lambda_n \in K$ mit $w = \lambda_1 v_1 + ... + \lambda_n v_n$. Daher ist $(-1) \cdot w + \lambda_1 v_1 + ... + \lambda_n v_n = \mathbf{0}$ mit $(0, ..., 0) \neq (-1, \lambda_1, ..., \lambda_n) \in K^{n+1}$, also $\{w, v_1, ..., v_n\}$ und damit $\mathcal{B}$ linear abhängig.

b)⇒ c) (durch ¬ c) ⇒ ¬ b)): Ist $V = \operatorname{lin}(\mathcal{B})$ und gibt es zu einem $v \in V$ zwei verschiedene Darstellungen als Linearkombination

$$v = \lambda_1 v_1 + ... + \lambda_n v_n = \mu_1 w_1 + ... + \mu_m w_m$$

mit $\lambda_1, ..., \lambda_n, \mu_1, ..., \mu_m \in K\backslash\{0\}$ und $v_1, ..., v_n, w_1, ..., w_m \in \mathcal{B}$, so dürfen wir (nach Einfügen von Vorfaktoren $= 0$ und eventuellem Umnumerieren) annehmen, daß

$$v = \tilde{\lambda}_1\tilde{v}_1 + ... + \tilde{\lambda}_N\tilde{v}_N = \tilde{\mu}_1\tilde{v}_1 + ...\tilde{\mu}_N\tilde{v}_N$$

mit $\tilde{\lambda}_1, ..., \tilde{\lambda}_N, \tilde{\mu}_1, ..., \tilde{\mu}_N \in K$, $\tilde{\lambda}_1 \neq \tilde{\mu}_1$ und $\tilde{v}_1, ..., \tilde{v}_N \in \mathcal{B}$. Auflösen nach $\tilde{v}_1$ ergibt

$$\tilde{v}_1 = \frac{\tilde{\mu}_2 - \tilde{\lambda}_2}{\tilde{\lambda}_1 - \tilde{\mu}_1}\tilde{v}_2 + ... + \frac{\tilde{\mu}_n - \tilde{\lambda}_n}{\tilde{\lambda}_1 - \tilde{\mu}_1}\tilde{v}_n,$$

also $\text{lin}(\mathcal{B}) = \text{lin}(\mathcal{B}\backslash\{\tilde{v}_1\})$.

c)$\Rightarrow$ d) Gilt c), und sind $v_1, ..., v_n \in \mathcal{B}$ paarweise verschieden so gibt es wegen $\mathbf{0} = 0v_1 + ... + 0v_n$ und der in c) vorausgesetzten Eindeutigkeit kein $(\lambda_1, ..., \lambda_n) \in K^n\backslash\{(0, ..., 0)\}$ mit $\lambda_1 v_1 + ... + \lambda_n v_n = \mathbf{0}$. Daher ist jede endliche Teilmenge $\{v_1, ..., v_n\}$ von $\mathcal{B}$ und damit $\mathcal{B}$ linear unabhängig. Ist nun $v \in V$ und $v = \lambda_1 v_1 + ... + \lambda_n v_n$ die nach c) eindeutige Darstellung von $v$ als Linearkombination mit $\lambda_1, ..., \lambda_n \in K\backslash\{0\}$, dann ist $\lambda_1 v_1 + ... + \lambda_n v_n + (-1)v = \mathbf{0}$, also $\{v_1, ..., v_n, v\}$ linear abhängig.

d)$\Rightarrow$ a) Gilt d), so gibt es für alle $v \in V$ Basiselemente $v_1, ..., v_n \in \mathcal{B}$ und $\lambda_1, ..., \lambda_n, \lambda \in K$, $\lambda \neq 0$ mit $\lambda_1 v_1 + ... + \lambda_n v_n + \lambda v = \mathbf{0}$, also ist $v = -\frac{\lambda_1}{\lambda}v_1 - ... - \frac{\lambda_n}{\lambda}v_n \in \text{lin}(\mathcal{B})$ und so $V = \text{lin}(\mathcal{B})$.

$\square$

**Beispiel 7.1.5.** Es sei

$$V := \left\{ \begin{array}{rcl} f : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & a_n z^n + ... + a_1 z + a_0 \end{array} : n \in \mathbb{N}_0, a_n, ..., a_1, a_0 \in \mathbb{C} \right\} \tag{7.1.2}$$

der $\mathbb{C}-$Vektorraum aller komplexen Polynomfunktionen. Dann bilden die **Monome**

$$\begin{array}{rcl} \tau_n : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & z^n \end{array}, n \in \mathbb{N}_0$$

eine Basis von $V$, denn offenbar ist $V = \text{lin}(\{\tau_n : n \in \mathbb{N}_0\})$. Sind nun $n_1, ..., n_k \in \mathbb{N}_0$ paarweise verschieden und $\lambda_1, ..., \lambda_k \in \mathbb{C}$ mit

$$\lambda_1 \tau_{n_1} + ... + \lambda_k \tau_{n_k} = \mathbf{0}$$

dh. für jedes $z \in \mathbb{C}$ ist

$$\lambda_1 z^{n_1} + ... + \lambda_k z^{n_k} = 0. \tag{7.1.3}$$

Gibt es nun ein $l \in \{1, ..., k\}$ mit $\lambda_l \neq 0$, dann ist

$$N := \max\{n_l : l = 1, ..., k, \lambda_l \neq 0\} \in \mathbb{N}_0$$

und somit $\lambda_1 \tau_{n_1} + ... + \lambda_k \tau_{n_k}$ eine Polynomfunktion vom Grad $N$. Diese hat höchstens $N$ Nullstellen [1] in $\mathbb{C}$ gezählt mit Vielfachheiten – wie man mit Polynomdivision sieht – aber dies steht im Widerspruch zu (7.1.3). Daher ist $\{\tau_n : n \in \mathbb{N}_0\}$ linear unabhängig in $V$ und damit eine Basis von $V$.

[1] nach dem Fundamentalsatz der Algebra genau $N$ Nullstellen in $\mathbb{C}$ (gezählt mit Vielfachheiten)

## 7.2 Endlichdimensionale Vektorräume

Wir betrachten im Folgenden immer paarweise verschiedene Vektoren $v_1, ..., v_n$.

**Lemma 7.2.1** (Austauschlemma). *Es sei $V$ ein $K$-Vektorraum und $\mathcal{B} = \{v_1, ..., v_n\}$ sei eine Basis von $V$, und $v = \lambda_1 v_1 + ... + \lambda_n v_n$. Ist $k \in \{1, ..., n\}$ mit $\lambda_k \neq 0$, dann ist $\mathcal{B}' = \{v_1, ..., v_{k-1}, v, v_{k+1}, ..., v_n\}$ wieder eine Basis von $V$.*

*Beweis.* Ist $w \in V$, so gibt es $\mu_1, ..., \mu_n \in K$ mit $w = \mu_1 v_1 + ... + \mu_n v_n$ und wegen $\lambda_k \neq 0$ ist

$$v_k = \frac{1}{\lambda_k} v - \frac{\lambda_1}{\lambda_k} v_1 - ... - \frac{\lambda_{k-1}}{\lambda_k} v_{k-1} - \frac{\lambda_{k+1}}{\lambda_k} v_{k+1} - ... - \frac{\lambda_n}{\lambda_k} v_n,$$

also

$$w = \frac{\mu_k}{\lambda_k} v + (\mu_1 - \frac{\mu_k \lambda_1}{\lambda_k}) v_1 + ... + (\mu_{k-1} - \frac{\mu_k \lambda_{k-1}}{\lambda_k}) v_{k-1} + (\mu_{k+1} - \frac{\mu_k \lambda_{k+1}}{\lambda_k}) v_{k+1} + ... + (\mu_n - \frac{\mu_k \lambda_n}{\lambda_k}) v_n$$

und damit $\operatorname{lin}(\mathcal{B}') = V$.

Ist nun $\mu_1 v_1 + ... + \mu_{k-1} v_{k-1} + \mu_k v + \mu_{k+1} v_{k+1} + ... + \mu_n v_n = \mathbf{0}$, so ergibt sich durch Einsetzen von $v = \lambda_1 v_1 + ... + \lambda_n v_n$:

$$(\mu_1 + \mu_k \lambda_1) v_1 + ... + (\mu_{k-1} + \mu_k \lambda_{k-1}) v_{k-1} + \mu_k \lambda_k v_k + (\mu_{k+1} + \mu_k \lambda_{k+1}) v_{k+1} + ... + (\mu_n + \mu_k \lambda_n) v_n = \mathbf{0}$$

Da $\mathcal{B}$ linear unabhängig ist, folgt daraus

$$\mu_1 + \mu_k \lambda_1 = ... = \mu_{k-1} + \mu_k \lambda_{k-1} = \mu_k \lambda_k v_k = \mu_{k+1} + \mu_k \lambda_{k+1} = ... = \mu_n + \mu_k \lambda_n = 0$$

Nach Voraussetzung ist $\lambda_k \neq 0$, also $\mu_k = 0$ und dann $\mu_1 = ... = \mu_n = 0$, also $\mathcal{B}'$ linear unabhängig. □

**Satz 7.2.2** (Basisaustauschsatz). *Es sei $V$ ein $K$-Vektorraum und $\mathcal{B} = \{v_1, ..., v_n\}$ eine Basis von $V$. $w_1, ..., w_m$ seien linear unabhängig. Dann ist $m \leq n$ und es gibt Indizes $i_1, ..., i_m \in \{1, ..., n\}$, so daß man durch den Austausch von $v_{i_k}$ durch $w_k$ für $k = 1, ..., m$ aus $\mathcal{B}$ wieder eine Basis von $V$ erhält.*

*Beweis.* durch Induktion nach $m$.

- Induktionsanfang $m = 1$:
  Ist $w_1$ linear unabhängig, so ist $w_1 \neq \mathbf{0}$, also gibt es für $w_1 = \lambda_1 v_1 + ... + \lambda_n v_n$ mindestens ein $k \in \{1, ..., n\}$ mit $\lambda_k \neq 0$ und nach dem Austauschlemma 7.2.1 ist dann $\{v_1, ..., v_{k-1}, w_1, v_{k+1}, ..., v_n\}$ eine Basis von $V$. Die Behauptung $m \leq n$ ist für $m, n \in \mathbb{N}$ und $m = 1$ erfüllt.

- Induktionsschritt: $m - 1 \to m$, $m \geq 2$:
  Auch $\{w_1, ..., w_{m-1}\}$ ist linear unabhängig, daher folgt aus der Induktionsannahme (nach geeigneter Umnummerierung) daß $\{w_1, ..., w_{m-1}, v_m, ..., v_n\}$ Basis von $V$ ist. Nach Induktionsannahme gilt ebenfalls $m - 1 \leq n$ und wir müssen zuerst den Fall $m - 1 = n$ ausschließen. In diesem Fall wäre aber schon $\{w_1, ..., w_{m-1}\}$ eine Basis von $V$, was Satz 7.1.4 widerspricht, denn danach ist $\{w_1, ..., w_m\} = \{w_1, ..., w_{m-1}\} \cup \{w_m\}$ linear abhängig. Schreibe nun $w_m = \lambda_1 w_1 + ... + \lambda_{m-1} w_{m-1} + \lambda_m v_m + ... + \lambda_n v_n$ mit $\lambda_1, ..., \lambda_n \in K$. Es gibt $k \in \{m, ..., n\}$ mit $\lambda_k \neq 0$ - denn sonst sind $w_1, ..., w_m$ linear abhängig. Daher läßt sich $v_k$ durch $w_m$ nach dem Austauschlemma ersetzen. □

**Korollar 7.2.3.** *Je zwei Basen eines endlich dimensionalen $K$-Vektorraums $V$ haben gleich viele Elemente.*

*Beweis.* Sind $\{v_1, ..., v_n\}$ und $\{w_1, ..., w_m\}$ zwei Basen von $V$, so kann man zweimal den Basisaustauschsatz anwenden:

- einmal für die Basis $\{w_1, ..., w_m\}$ und die linear unabhängigen Vektoren $v_1, ..., v_n$, was $n \leq m$ ergibt
- einmal für die Basis $\{v_1, ..., v_n\}$ und die linear unabhängigen Vektoren $w_1, ..., w_m$, was $m \leq n$ gibt.

□

**Definition 7.2.4.** *Ist $V$ ein endlich dimensionaler $K$-Vektorraum und $\{v_1, ..., v_n\}$ eine Basis von $V$, so heißt*

$$n =: \dim_K(V)$$

*die Dimension von $V$ (über $K$).*

**Korollar 7.2.5** (Basisergänzungssatz). *Es sei $V$ ein $K$-Vektorraum mit $n = \dim_K V < \infty$ und $v_1, ..., v_m$ seien linear unabhängig in $V$. Dann kann man $v_{m+1}, ..., v_n \in V$ finden, so daß $\mathcal{B} = \{v_1, ..., v_n\}$ eine Basis von $V$ ist.*

*Beweis.* Es sei $\{w_1, ..., w_n\}$ eine Basis von $V$. Anwenden des Basisaustauschsatzes auf die linear unabhängigen Vektoren $v_1, ..., v_m$ ergibt (nach eventuellem Umnummerieren) eine Basis $\{v_1, ..., v_m, w_{m+1}, ..., w_n\}$ von $V$. □

**Korollar 7.2.6.** *Ist $V$ ein $K$-Vektorraum, $n = \dim_K V$ und $W \subseteq V$ ein Untervektorraum, dann ist $\dim_K W \leq n$. Gilt in diesem Fall $\dim_K W = n$, so folgt $W = V$.*

*Beweis.* Es sei $\{w_1, ..., w_m\}$ eine Basis von $W$, also insbesondere linear unabhängig in $V$. Dann läßt sich nach dem Basisergänzungssatz 7.2.5 eine Basis $\{w_1, ..., w_m, v_{m+1}, ..., v_n\}$ von $V$ bekommen. Insbesondere ist $m \leq n$ und für $m = n$ ist $W = \text{lin}(\{w_1, ..., w_m\}) = V$. □

**Definition 7.2.7.** *Es sei $V$ ein $K$-Vektorraum und $V_1, ..., V_m \subseteq V$ seien Untervektorräume, dann heißt*

$$V_1 + ... + V_m := \{v \in V : \text{ Es gibt } v_1 \in V_1, ..., v_m \in V_m \text{ mit } v = v_1 + ... + v_m\} \quad (7.2.1)$$

*die* **Summe** *der Untervektorräume $V_1, ..., V_m$.*

**Lemma 7.2.8.** *Sind $V_1, ..., V_m$ Untervektorräume von $V$, dann ist*

$$V_1 + ... + V_m = \text{lin}(V_1 \cup ... \cup V_m) \quad (7.2.2)$$

*der kleinste Untervektorraum, der $V_1 \cup ... \cup V_m$ enthält.*

*Beweis.* Es bleibt nur noch die Gleichheit (7.2.2) zu zeigen; die restlichen Aussagen folgen aus Lemma 3.3.8, wo

$$\mathrm{lin}(X) = \bigcap_{\substack{W \subseteq V \text{ Untervektorraum} \\ X \subseteq W}} W$$

für jedes $\emptyset \neq X \subseteq V$ gezeigt wurde. Da $V_1, ..., V_m$ Untervektorräume sind, folgt aus $\lambda_j \in K$ und $w_j \in V_j$ auch $v_j = \lambda_j w_j \in V_j$ für $j = 1, ..., m$. Beachtet man noch, daß in der allgemeinst möglichen Linearkombination $\sum_{k=1}^{n} \alpha_k u_k$ mit $u_1, ..., u_n \in V_1 \cup ... \cup V_m$, die in $\mathrm{lin}(V_1 \cup ... \cup V_m)$ enthalten ist, es zu jedem $u_k$ ein $j \in \{1, ..., m\}$ gibt, so daß $u_k \in V_j$ gilt, so finden wir (indem wir zuerst alle Vektoren dieser Linearkombination aus $V_1$, dann die aus $V_2$, etc. zusammenfassen) $\lambda_1, ..., \lambda_m \in K$ und $w_1 \in V_1, ..., w_m \in V_m$ mit $\sum_{k=1}^{n} \alpha_k u_k = \sum_{j=1}^{m} \lambda_j w_j$ und damit ist

$$\begin{aligned}
\mathrm{lin}(V_1 \cup ... \cup V_m) &= \{v \in V : v = \sum_{j=1}^{m} \lambda_j w_j, \lambda_1, ..., \lambda_m \in K, w_1 \in V_1, ..., w_m \in V_m\} \\
&= \{v \in V : v = \sum_{j=1}^{m} v_j, v_1 \in V_1, ..., v_m \in V_m\} = V_1 + ... + V_m
\end{aligned}$$

□

**Satz 7.2.9.** *Es sei $V$ ein endlichdimensionaler $K$-Vektorraum und $U, W \subseteq V$ Untervektorräume, dann ist*

$$\dim_K(U + W) + \dim_K(U \cap W) = \dim_K(U) + \dim_K(W) \tag{7.2.3}$$

*Beweis.* Es sei

$$\mathcal{B}_{U \cap W} = \{f_1, ..., f_k\}$$

eine Basis von $U \cap W$. Ergänze diese mittels des Basisergänzungssatzes zu einer Basis

$$\mathcal{B}_U = \{f_1, ..., f_k, g_1, ..., g_l\}$$

von $U$ und zu einer Basis

$$\mathcal{B}_W = \{f_1, ..., f_k, h_1, ..., h_m\}$$

von $W$. Betrachte nun

$$\mathcal{B} := \{f_1, ..., f_k, g_1, ..., g_l, h_1, ..., h_m\},$$

dann ist offenbar $\mathrm{lin}(\mathcal{B}) = U + W$. Sind nun $\alpha_1, ..., \alpha_k, \beta_1, ..., \beta_l, \gamma_1, ..., \gamma_m \in K$ mit

$$\sum_{j=1}^{k} \alpha_j f_j + \sum_{j=1}^{l} \beta_j g_j + \sum_{j=1}^{m} \gamma_j h_j = \mathbf{0}, \tag{7.2.4}$$

dann gilt:

$$\underbrace{\sum_{j=1}^{k} \alpha_j f_j + \sum_{j=1}^{l} \beta_j g_j}_{\in U} = \underbrace{-\sum_{j=1}^{m} \gamma_j h_j}_{\in W}.$$

Damit ist $-\sum_{j=1}^{m} \gamma_j h_j \in U \cap W$ und da $\mathcal{B}_{U\cap W}$ eine Basis von $U \cap W$ ist, gibt es $\lambda_1, ..., \lambda_k \in K$ mit $-\sum_{j=1}^{m} \gamma_j h_j = \sum_{j=1}^{k} \lambda_j f_j$, also

$$\sum_{j=1}^{m} \gamma_j h_j + \sum_{j=1}^{k} \lambda_j f_j = \mathbf{0}. \tag{7.2.5}$$

Da $\mathcal{B}_W$ als Basis von $W$ linear unabhängig ist folgt $\lambda_1 = ... = \lambda_k = \gamma_1 = ... = \gamma_m = 0$ aus (7.2.5). Damit reduziert sich (7.2.4) auf

$$\sum_{j=1}^{k} \alpha_j f_j + \sum_{j=1}^{l} \beta_j g_j = \mathbf{0} \tag{7.2.6}$$

woraus $\alpha_1 = ... = \alpha_k = \beta_1 = ... = \beta_l = 0$ folgt, da auch $\mathcal{B}_U$ linear unabhängig ist. Dies zeigt, daß $\mathcal{B}$ linear unabhängig und daher Basis von $U + W$ ist, was auch die Dimensionsformel (7.2.3) zeigt. □

**Definition 7.2.10.** *Es sei $V$ ein $K$-Vektorraum und $U_1, ..., U_m \subseteq V$ seien Untervektorräume, dann heißt $W = U_1 + ... + U_m$ die* **direkte Summe** *der Untervektorräume $U_1, ..., U_m$, wenn jede Zerlegung von $w \in W$ der Form $w = u_1 + ... + u_m$ mit $u_1 \in U_1, ..., u_m \in U_m$ eindeutig ist. In diesem Fall schreibe $W = \bigoplus_{j=1}^{m} U_j$ für die direkte Summe.*

**Beispiel 7.2.11.**

a) Ist $V$ ein $K$-Vektorraum und $\mathcal{B} = \{v_1, ..., v_n\}$ eine Basis von $V$, dann ist $V = \bigoplus_{j=1}^{n} \operatorname{lin}(\{v_j\})$, denn da $\mathcal{B}$ ein Erzeugendensystem ist, gilt $\operatorname{lin}(\mathcal{B}) = V$ und nach Satz 7.1.4 sind für jedes $v \in V$ die Koeffizienten $\lambda_1, ..., \lambda_n \in K$ der Linearkombination $v = \lambda_1 v_1 + ... + \lambda_n v_n$ eindeutig bestimmt.

b) Im $\mathbb{R}-$Vektorraum $V = \mathbb{R}^3$ bilden $U_1 := \operatorname{lin}\left(\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right\}\right)$ und $U_2 := \operatorname{lin}\left(\left\{\begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right\}\right)$ keine direkte Summe, denn zB.

$$\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}1\\1\\0\end{pmatrix} + \begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} + \begin{pmatrix}1\\1\\1\end{pmatrix}$$

sind zwei verschiedene Zerlegungen mit Summanden in $U_1$ und $U_2$.

**Satz 7.2.12.** *Es sei $V$ ein $K$-Vektorraum und $U_1, ..., U_m \subseteq V$ Untervektorräume, dann sind äquivalent:*

*a)* $W = \bigoplus_{j=1}^{m} U_j$

*b) Es ist* $W = U_1 + ... + U_m$ *und für alle* $j = 1, ..., m$ *und* $V_j := \sum_{\substack{k=1 \\ k \neq j}}^{m} U_k$ *gilt:* $U_j \cap V_j = \{\mathbf{0}\}$.

*In diesem Fall ist*

$$\dim_K \left( \bigoplus_{j=1}^{m} U_j \right) = \sum_{j=1}^{m} \dim_K(U_j). \tag{7.2.7}$$

*Beweis.*

a)⇒b) $W = \bigoplus_{j=1}^{m} U_j$ bedeutet per Definition, daß jedes $w \in W$ eine eindeutige Zerlegung der Form $w = u_1 + ... + u_m$ mit $u_1 \in U_1$,...,$u_m \in U_m$ besitzt. Damit ist insbesondere $W = U_1 + ... + U_m$. Es sei nun $j \in \{1, ..., m\}$ und $w \in U_j \cap V_j$. Wenn $w \neq \mathbf{0}$ ist, so finden wir (wegen $w \in V_j$) Vektoren $u_1 \in U_1$,...,$u_{j-1} \in U_{j-1}$,$u_{j+1} \in U_{j+1}$,...,$u_m \in U_m$ mit $w = u_1 + ... + u_{j-1} + u_{j+1} + ... + u_m$. Wegen $w \in U_j$ ist $-w \in U_j$, also $\mathbf{0} = w + (-w) = u_1 + ... + u_{j-1} - w + u_{j+1} + ... + u_m$ eine von $\mathbf{0} = \mathbf{0} + ... + \mathbf{0}$ verschiedene Zerlegung im Widerspruch zu $W = \bigoplus_{j=1}^{m} U_j$.

b)⇒a) Gilt b) und ist $W$ nicht die direkte Summe, dann gibt es für ein $w \in W$ verschiedene Zerlegungen $w = u_1 + ... + u_m = \tilde{u}_1 + ... + \tilde{u}_m$ mit $u_1, \tilde{u}_1 \in U_1$,..., $u_m, \tilde{u}_m \in U_m$ und ein $j \in \{1, ..., m\}$ mit $u_j \neq \tilde{u}_j$. Dann ist

$$U_j \ni \tilde{u}_j - u_j = \sum_{\substack{k=1 \\ k \neq j}}^{m} (u_k - \tilde{u}_k) \in V_j$$

also $\mathbf{0} \neq \tilde{u}_j - u_j \in U_j \cap V_j$ im Widerspruch zur Voraussetzung in b).

Beweis der Dimensionsformel (7.2.7) durch Induktion nach $m$:

- Induktionsanfang $m = 1$ ist $\dim_K(U_1) = \dim_K(U_1)$.
- Induktionsschritt $m \to m + 1$:
  Ist $W_{m+1} := \bigoplus_{j=1}^{m+1} U_j = W_m \oplus U_{m+1} = \left( \bigoplus_{j=1}^{m} U_j \right) \oplus U_{m+1}$, so gilt $W_{m+1} = W_m + U_{m+1}$ und $W_m \cap U_{m+1} = \{\mathbf{0}\}$ nach b). Anwenden der Dimensionsformel aus Satz 7.2.9 und die Verwendung der Induktionsvoraussetzung auf $W_m$ zeigt:

$$\begin{aligned} \dim_K(W_{m+1}) &= \dim_K(W_{m+1}) + \dim_K(W_m \cap U_{m+1}) = \dim_K(W_m) + \dim_K U_{m+1} \\ &= \dim_K \left( \bigoplus_{j=1}^{m} U_j \right) + \dim_K(U_{m+1}) = \sum_{j=1}^{m+1} \dim_K(U_j) \end{aligned}$$

□

# Kapitel 8

# Lineare Abbildungen und Matrizen

## 8.1 Lineare Abbildungen

**Definition 8.1.1.** *Es seien $V$ und $W$ $K$-Vektorräume. Eine Abbildung $F : V \to W$ heißt* $K$**-linear***, wenn für alle $u, v \in V$ und $\lambda \in K$ gilt:*

$$\begin{aligned} F(u+v) &= F(u) + F(v) \qquad (8.1.1) \\ F(\lambda v) &= \lambda F(v) \qquad (8.1.2) \end{aligned}$$

*Eine oft verwendete Schreibweise ist*

$$\mathrm{Hom}_K(V, W) := \{F : V \to W : F \text{ ist } K\text{-linear}\}. \qquad (8.1.3)$$

*Eine $K$-lineare Abbildung $F : V \to W$ heißt ein*

- **Isomorphismus***, wenn $F$ bijektiv ist.*
- **Endomorphismus***, wenn $V = W$ ist.*
- **Automorphismus***, wenn $V = W$ und $F$ bijektiv ist.*

**Beispiel 8.1.2.** Ist $a \in \mathbb{C}$, so ist $\begin{array}{rcl} f : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & az \end{array}$ eine $\mathbb{C}$-lineare Abbildung – das folgt aus den Rechenregeln für komplexe Zahlen. Ist $b \neq 0$, so ist $\begin{array}{rcl} g : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & az + b \end{array}$ keine $\mathbb{C}$-lineare Abbildung – das sieht man etwa an $g(0+0) = b \neq g(0) + g(0) = b + b$.

**Satz 8.1.3.** *Ist $F : V \to W$ eine $K$-lineare Abbildung, dann gilt:*

*a)* $F(\mathbf{0}) = \mathbf{0}$

*b) Für $\lambda_1, ..., \lambda_n \in K$ und $v_1, ..., v_n \in V$ gilt*

$$F(\lambda_1 v_1 + ... + \lambda_n v_n) = \lambda_1 F(v_1) + ... + \lambda_n F(v_n)$$

*c) Ist $(v_i)_{i \in I}$ linear abhängig in $V$, so ist auch $(F(v_i))_{i \in I}$ linear abhängig in $W$.*

*d) Ist $U \subseteq V$ ein Untervektorraum, dann ist auch $F(U) \subseteq W$ ein Untervektorraum.*

e) *Ist $Y \subseteq W$ ein Untervektorraum, dann ist auch das Urbild $F^{-1}(Y) \subseteq V$ ein Untervektorraum. Insbesondere ist* $\mathrm{Kern} F := F^{-1}(\{0\})$ *ein Untervektorraum von $V$.*

f) *Ist $F : V \to W$ ein Isomorphismus, dann ist auch die Umkehrabbildung $F^{-1} : W \to V$ eine $K$-lineare Abbildung.*

*Beweis.*

a) Nach Lemma 3.3.3 gilt $0 \cdot v = \mathbf{0}$ für alle $v \in V$ (bzw. $0 \cdot w = \mathbf{0}$ für alle $w \in W$); eingesetzt in (8.1.2) ergibt dies: $F(\mathbf{0}) = F(0 \cdot \mathbf{0}) = 0 \cdot F(\mathbf{0}) = \mathbf{0}$.

b) Durch Induktion nach $n$. (8.1.2) zeigt den Induktionsanfang für $n = 1$. Für den Induktionsschluß $n-1 \to n$ bemerken wir, daß $u := \lambda_1 v_1 + ... + \lambda_{n-1} v_{n-1} \in V$, also nach (8.1.1) und (8.1.2) $F(\lambda_1 v_1 + ... + \lambda_n v_n) = F(u + \lambda_n v_n) = F(u) + \lambda_n F(v_n)$ gilt. Anwenden der Induktionsannahme auf $F(u)$ ergibt dann

$$F(\lambda_1 v_1 + ... + \lambda_n v_n) = \lambda_1 F(v_1) + ... + \lambda_{n-1} F(v_{n-1}) + \lambda_n F(v_n).$$

c) Sind $(v_i)_{i \in I}$ linear abhängig, so gibt es $n \in \mathbb{N}$, $(\lambda_1, ..., \lambda_n) \in K^n \backslash \{(0, ..., 0)\}$ und $v_{i_1}, ..., v_{i_n}$ mit $\lambda_1 v_{i_1} + ... + \lambda_n v_{i_n} = \mathbf{0}$. Mit Teil a) und b) folgt dann

$$\mathbf{0} = F(\mathbf{0}) = F(\lambda_1 v_{i_1} + ... + \lambda_n v_{i_n}) = \lambda_1 F(v_{i_1}) + ... + \lambda_n F(v_{i_n}),$$

also ist auch $(F(v_i))_{i \in I}$ linear abhängig.

d) Da $U \subseteq V$ ein Untervektorraum ist, gilt $\mathbf{0} \in U$ und $\lambda_1 u_1 + \lambda_2 u_2 \in U$ für $u_1, u_2 \in U$ und $\lambda_1, \lambda_2 \in K$. Damit ist auch $\mathbf{0} = F(\mathbf{0}) \in F(U)$ und zu $w_1 = F(u_1) \in F(U)$ und $w_2 = F(u_2) \in F(U)$ ist auch $\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 F(u_1) + \lambda_2 F(u_2) = F(\lambda_1 u_1 + \lambda_2 u_2) \in F(U)$, also $F(U)$ nach Definition 3.3.4 ein Untervektorraum.

e) Wegen $\mathbf{0} = F(\mathbf{0})$ ist $\mathbf{0} \in F^{-1}(Y)$. Sind nun $u_1, u_2 \in F^{-1}(Y)$, so gibt es $w_1, w_2 \in Y$ mit $F(u_1) = w_1$ und $F(u_2) = w_2$, also ist $F(\lambda_1 u_1 + \lambda_2 u_2) = \lambda_1 w_1 + \lambda_2 w_2 \in Y$ und damit $\lambda_1 u_1 + \lambda_2 u_2 \in F^{-1}(Y)$, dh. $F^{-1}(Y)$ ein Untervektorraum.

f) Ist $F : V \to W$ ein Isomorphismus, so ist $F$ bijektiv und die Umkehrabbildung $F^{-1} : W \to V$ existiert. Zu $w_1, w_2 \in W$ gibt es jeweils genau ein $v_1, v_2 \in V$ mit $w_1 = F(v_1)$ und $w_2 = F(v_2)$ und daraus folgt $w_1 + \lambda w_2 = F(v_1 + \lambda v_2)$ für alle $\lambda \in K$. Durch Anwenden von $F^{-1}$ ergibt sich: $F^{-1}(w_1 + \lambda w_2) = v_1 + \lambda v_2 = F^{-1}(w_1) + \lambda F^{-1}(w_2)$, dh. $F^{-1}$ ist linear. □

**Lemma 8.1.4.** *Es seien $U, V, W$ $K$-Vektorräume, $G : U \to V$ und $F : V \to W$ seien $K$-linear, dann ist auch $F \circ G : U \to W$ $K$-linear.*

*Beweis.* Sind $u_1, u_2 \in U$ und $\lambda \in K$, dann ist $(F \circ G)(u_1 + \lambda u_2) = F(G(u_1 + \lambda u_2)) = F(G(u_1)) + \lambda F(G(u_2)) = (F \circ G)(u_1) + \lambda (F \circ G)(u_2)$. □

**Lemma 8.1.5.** *Es sei $F : V \to W$ $K$-linear, dann ist $F$ genau dann injektiv, wenn* $\mathrm{Kern} F = \{\mathbf{0}\}$ *ist. Sind in diesem Fall $v_1, ..., v_n$ linear unabhängig in $V$, dann ist auch $\{F(v_1), ..., F(v_n)\}$ linear unabhängig in $W$.*

*Beweis.* Wegen $F(\mathbf{0}) = \mathbf{0}$ gilt $\{\mathbf{0}\} \subseteq \mathrm{Kern}F = F^{-1}(\{\mathbf{0}\})$. Ist $F$ injektiv, so hat $F^{-1}(\{\mathbf{0}\})$ höchstens ein Element, damit folgt $\mathrm{Kern}F = \{\mathbf{0}\}$. Ist umgekehrt $\mathrm{Kern}F = \{\mathbf{0}\}$ und sind $v_1, v_2 \in V$ mit $F(v_1) = F(v_2)$, dann gilt $\mathbf{0} = F(v_1) - F(v_2) = F(v_1 - v_2)$, dh. $v_1 - v_2 \in KernF = \{\mathbf{0}\}$ also $v_1 = v_2$, womit $F$ injektiv ist.

Ist nun $F$ injektiv und $v_1, ..., v_n \in V$ linear unabhängig und es seien $\lambda_1, ..., \lambda_n \in K$ mit $\lambda_1 F(v_1) + ... + \lambda_n F(v_n) = \mathbf{0}$. Dann ist $\lambda_1 F(v_1) + ... + \lambda_n F(v_n) = F(\lambda_1 v_1 + ... + \lambda_n v_n) = \mathbf{0}$, also $\lambda_1 v_1 + ... + \lambda_n v_n \in \mathrm{Kern}F = \{\mathbf{0}\}$ und aus der linearen Unabhängigkeit von $v_1, ..., v_n$ folgt $\lambda_1 = ... = \lambda_n = 0$ und damit sind auch $F(v_1), ..., F(v_n)$ linear unabhängig. □

Weitere Beispiele für lineare Abbildungen zwischen endlichdimensionalen Vektorräumen liefert der folgende Satz:

**Satz 8.1.6.** *Es seien $V, W$ endlichdimensionale $K$-Vektorräume, $\mathcal{B} = \{v_1, ..., v_n\}$ sei eine Basis von $V$ und $w_1, ..., w_n \in W$. Dann gibt es genau ein*

$$F : V \to W \in \mathrm{Hom}_K(V, W) \qquad \textit{mit} \qquad F(v_1) = w_1, ..., F(v_n) = w_n. \tag{8.1.4}$$

*$F$ ist also bereits durch die Bilder $F(v_1), ..., F(v_n)$ der Basisvektoren $v_1, ..., v_n$ eindeutig bestimmt.*
*Zusatz: Für diese $K$-lineare Abbildung $F$ gilt immer $F(V) = \mathrm{lin}(\{w_1, ..., w_n\})$. Ferner ist $F$ genau dann injektiv, wenn $\{w_1, ..., w_n\}$ linear unabhängig ist.*

*Beweis.* Da $\mathcal{B}$ eine Basis von $V$ ist, besitzt jedes $v \in V$ eine eindeutige Darstellung $v = \lambda_1 v_1 + ... + \lambda_n v_n$. Die beiden Forderungen $F$ linear und $F(v_1) = w_1$,...,$F(v_n) = w_n$ lassen damit nur

$$F(v) = F(\lambda_1 v_1 + ... + \lambda_n v_n) := \lambda_1 w_1 + ... + \lambda_n w_n \tag{8.1.5}$$

zu; daher gibt es höchstens ein $F$, das (8.1.4) erfüllt. Die durch (8.1.5) definierte Abbildung ist auch $K$-linear, wie man für $u = \mu_1 v_1 + ... + \mu_n v_n \in V$ und $\lambda \in K$ aus

$$\begin{aligned} F(u + \lambda v) &= F((\mu_1 + \lambda\lambda_1)v_1 + ... + (\mu_n + \lambda\lambda_n)v_n) \\ &= (\mu_1 + \lambda\lambda_1)w_1 + ... + (\mu_n + \lambda\lambda_n)w_n \\ &= \mu_1 w_1 + ... + \mu_n w_n + \lambda(\lambda_1 w_1 + ... + \lambda_n w_n) = F(u) + \lambda F(v) \end{aligned}$$

sieht. Offenbar ist $F(V) \subseteq \mathrm{lin}(\{w_1, ..., w_n\})$ und für $w = \mu_1 w_1 + ... + \mu_n w_n \in \mathrm{lin}(\{w_1, ..., w_n\})$ gilt $w = F(\mu_1 v_1 + ... + \mu_n v_n)$, dh. $\mathrm{lin}(\{w_1, ..., w_n\}) \subseteq F(V)$.

„$\Rightarrow$“ Ist $\{w_1, ..., w_n\}$ linear abhängig, dann gibt es $(\lambda_1, ..., \lambda_n) \in K^n \backslash \{(0, ..., 0)\}$ mit $\lambda_1 w_1 + ... + \lambda_n w_n = \mathbf{0}$ und damit ist auch $F(\lambda_1 v_1 + ... + \lambda_n v_n) = \lambda_1 w_1 + ... + \lambda_n w_n = \mathbf{0}$, also $F$ nicht injektiv, da $\lambda_1 v_1 + ... + \lambda_n v_n \neq \mathbf{0}$ .

„$\Leftarrow$“ Sind $\{w_1, ..., w_n\}$ linear unabhängig, und $v = \lambda_1 v_1 + ... + \lambda_n v_n \in \mathrm{Kern}F$, dann ist $F(v) = \lambda_1 w_1 + ... + \lambda_n w_n = \mathbf{0}$ und wegen der linearen Unabhängigkeit von $w_1, .., w_n$ folgt $\lambda_1 = ... = \lambda_n = 0$. Folglich ist $\mathrm{Kern}F = \{\mathbf{0}\}$ und $F$ injektiv nach Lemma 8.1.5.

□

**Satz 8.1.7** (Dimensionsformel). *Es seien $V, W$ $K$-Vektorräume, $n = \dim_K V < \infty$ und $F : V \to W$ sei $K$-linear. Sind $\{v_1, ..., v_s\}$ eine Basis von $\mathrm{Kern}F$, $\{w_1, ..., w_r\}$ eine Basis von $F(V)$ und es seien $u_1 \in F^{-1}(\{w_1\}), ..., u_r \in F^{-1}(\{w_r\})$ gewählt. Dann ist $r + s = n$ also*

$$\dim_K V = \dim_K(\mathrm{Kern}F) + \dim_K(F(V)) \tag{8.1.6}$$

*und $\mathcal{A} = \{u_1, ..., u_r, v_1, ..., v_s\}$ ist eine Basis von $V$.*

*Beweis.* Es sei $v \in V$, $F(v) = \lambda_1 w_1 + ... + \lambda_r w_r$ und $u := \lambda_1 u_1 + ... + \lambda_r u_r \in V$. Wegen $F(u) = F(\lambda_1 u_1 + ... + \lambda_r u_r) = \lambda_1 F(u_1) + ... + \lambda_r F(u_r) = \lambda_1 w_1 + ... + \lambda_r w_r = F(v)$ ist $F(v - u) = \mathbf{0}$ oder $v - u \in \mathrm{Kern}F$. Es gibt also $\mu_1, ..., \mu_s \in K$ mit $v - u = \mu_1 v_1 + ... + \mu_s v_s$. Damit ist $v = \mu_1 v_1 + ... + \mu_s v_s + \lambda_1 u_1 + ... + \lambda_r u_r \in \mathrm{lin}(\mathcal{A})$ und daher $V = \mathrm{lin}(\mathcal{A})$. Ist nun

$$\mu_1 v_1 + ... + \mu_s v_s + \lambda_1 u_1 + ... + \lambda_r u_r = \mathbf{0}, \tag{8.1.7}$$

dann ist

$$\begin{aligned} \mathbf{0} &= F(\mu_1 v_1 + ... + \mu_s v_s + \lambda_1 u_1 + ... + \lambda_r u_r) \\ &= F(\mu_1 v_1 + ... + \mu_s v_s) + F(\lambda_1 u_1 + ... + \lambda_r u_r) = F(\lambda_1 u_1 + ... + \lambda_r u_r) \end{aligned}$$

da $\mu_1 v_1 + ... + \mu_s v_s \in \mathrm{Kern}F$. Da $F$ linear ist, gilt

$$\mathbf{0} = F(\lambda_1 u_1 + ... + \lambda_r u_r) = \lambda_1 F(u_1) + ... + \lambda_r F(u_r) = \lambda_1 w_1 + ... + \lambda_r w_r$$

und aus der linearen Unabhängigkeit von $w_1, ..., w_r$ folgt $\lambda_1 = ... = \lambda_r = 0$. Einsetzen in (8.1.7) ergibt $\mu_1 v_1 + ... + \mu_s v_s = \mathbf{0}$ und aus der linearen Unabhängigkeit von $\{v_1, ..., v_s\}$ folgt $\mu_1 = ... = \mu_s = 0$. Folglich ist $\mathcal{A}$ linear unabhängig, also wegen $V = \mathrm{lin}(\mathcal{A})$ eine Basis von $V$. □

**Korollar 8.1.8.** *Sind $V$ und $W$ endlichdimensionale $K$-Vektorräume, so gibt es genau dann einen Isomorphismus $f : V \to W$, wenn $\dim_K V = \dim_K W$ ist.*

*Beweis.*

„$\Rightarrow$“ Ist $f : V \to W$ ein Isomorphismus, so gilt $f(V) = W$ und $\mathrm{Kern}f = \{\mathbf{0}\}$. Aus der Dimensionsformel folgt $\dim_K V = \dim_K(\mathrm{Kern}f) + \dim_K(f(V)) = 0 + \dim_K W$.

„$\Leftarrow$“ Ist $n = \dim_K V = \dim_K W$, so gibt es eine Basis $\{v_1, ..., v_n\}$ von $V$ und eine Basis $\{w_1, ..., w_n\}$ von $W$. Die lineare Abbildung $F : V \to W$ mit $F(v_j) = w_j, j = 1, ..., n$ aus Satz 8.1.6 ist dann – nach dem Zusatz von Satz 8.1.6 – ein Isomorphismus.

□

## 8.2 Matrizen

**Definition 8.2.1.** *Sind $m, n \in \mathbb{N}$, so bezeichnen wir mit*

$$M(m \times n, K) := \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} : a_{ij} \in K \text{ für } 1 \le i \le m, 1 \le j \le n \right\}$$

*die Menge aller $m \times n$ Matrizen mit Einträgen aus dem Körper $K$. Die horizontalen Zahlenreihen heißen Zeilenvektoren der Matrix, die vertikalen Zahlenreihen heißen Spaltenvektoren der Matrix.*

**Beispiel 8.2.2.** Für $A = \begin{pmatrix} 5 & i \\ 12 & 3 \\ 7 & 0 \end{pmatrix} \in M(3 \times 2, \mathbb{C})$ ist $(5, i)$ der 1. Zeilenvektor von $A$ und $\begin{pmatrix} i \\ 3 \\ 0 \end{pmatrix}$ der 2. Spaltenvektor von $A$.

**Definition 8.2.3** (Addition und Skalarmultiplikation von Matrizen)**.**
*Sind* $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} \in M(m \times n, K)$, *so definiert*

$$A + B := \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix} \in M(m \times n, K)$$

*die Summe von $A$ und $B$. Für $\lambda \in K$ heißt*

$$\lambda \cdot A := \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}$$

*die Skalarmultiplikation von $\lambda$ und $A$.*

**Bemerkung 8.2.4.** Durch diese Addition $+$ und Skalarmultiplikation $\cdot$ wird $M(m \times n, K)$ zu einem Vektorraum über $K$. Die Nullmatrix $0 = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$ ist das neutrale Element von $(M(m \times n, K), +)$ und $-A := \begin{pmatrix} -a_{11} & \cdots & -a_{1n} \\ \vdots & & \vdots \\ -a_{m1} & \cdots & -a_{mn} \end{pmatrix}$ das inverse Element von $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ in $(M(m \times n, K), +)$. Der Beweis der weiteren Vektorraumeigenschaften von $M(m \times n, K)$ folgt aus den Rechenregeln in $K$ angewandt auf jeden Eintrag $a_{ij}$ in der Matrix $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$.

**Definition 8.2.5** (Multiplikation von Matrizen)**.** *Sind*
$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M(m \times n, K)$ *und* $B = \begin{pmatrix} b_{11} & \cdots & b_{1r} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nr} \end{pmatrix} \in M(n \times r, K)$,

*so wird das Produkt*

$$A \cdot B := \begin{pmatrix} c_{11} & \cdots & c_{1r} \\ \vdots & & \vdots \\ c_{m1} & \cdots & c_{mr} \end{pmatrix} \in M(m \times r, K)$$

*durch folgende Vorschrift für die Matrixelemente* $c_{ik}$

$$c_{ik} := a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk} \tag{8.2.1}$$

*definiert.*

**Beachte:** Um $A \cdot B$ überhaupt bilden zu können, muß die Anzahl der Spalten von $A$ mit der Anzahl der Zeilen von $B$ übereinstimmen. $A \cdot B$ hat dann so viele Zeilen wie $A$ und so viele Spalten wie $B$.

**Lemma 8.2.6.** *Sind* $A, A' \in M(m \times n, K)$, $B, B' \in M(n \times r, K)$, $C \in M(r \times s, K)$ *und* $\lambda \in K$, *so gelten:*

*a) Distributivgesetze:*

$$\begin{aligned} A \cdot (B + B') &= A \cdot B + A \cdot B' \\ (A + A') \cdot B &= A \cdot B + A' \cdot B \end{aligned}$$

*b)* $A \cdot (\lambda B) = (\lambda A) \cdot B = \lambda(A \cdot B)$

*c) Assoziativgesetz:* $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

*Beweis.* Zum Beweis von c) schreibe $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$, $B = (b_{jk})_{\substack{j=1,\dots,n \\ k=1,\dots,r}}$ $C = (c_{kl})_{\substack{k=1,\dots,r \\ l=1,\dots,s}}$ und die Matrixprodukte aus zwei Matrizen als $A \cdot B = (\alpha_{ik})_{\substack{i=1,\dots,m \\ k=1,\dots,r}}$, $B \cdot C = (\beta_{jl})_{\substack{j=1,\dots,n \\ l=1,\dots,s}}$ und die verschieden geklammerten Produkte aller drei Matrizen $(A \cdot B) \cdot C = (d_{il})_{\substack{i=1,\dots,m \\ l=1,\dots,s}}$ und $A \cdot (B \cdot C) = (f_{il})_{\substack{i=1,\dots,m \\ l=1,\dots,s}}$. Die Einträge $\alpha_{ik}$, $\beta_{jl}$ ergeben sich als $\alpha_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$ und $\beta_{jl} = \sum_{k=1}^{r} b_{jk}c_{kl}$, also folgt

$$d_{il} = \sum_{k=1}^{r} \alpha_{ik}c_{kl} = \sum_{k=1}^{r} (\sum_{j=1}^{n} a_{ij}b_{jk})c_{kl}$$

$$f_{il} = \sum_{j=1}^{n} a_{ij}\beta_{jl} = \sum_{j=1}^{n} a_{ij} \sum_{k=1}^{r} b_{jk}c_{kl} = d_{il}.$$

a) und b) können wir für jeden Eintrag einer Zeile und Spalte einzeln überprüfen. $\square$

**Bemerkung 8.2.7.** Die Matrixmultiplikation läßt sich durch folgendes Schema einprägen:

$$\begin{array}{c|c}
 & \begin{pmatrix} b_{11} & \cdots & b_{1k} & \cdots & b_{1r} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nk} & \cdots & b_{nr} \end{pmatrix} \\ \hline
\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} & \begin{pmatrix} c_{11} & \cdots & \cdot & \cdots & c_{1r} \\ \vdots & & \vdots & & \vdots \\ \cdot & & c_{ik} & & \cdot \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \cdots & \cdot & \cdots & c_{mr} \end{pmatrix}
\end{array}$$

**Merkregel:** "k-te Spalte der zweiten Matrix auf die i-te Zeile der ersten Matrix legen, alle Produkte $a_{i1}b_{1k}, ..., a_{in}b_{nk}$ bilden und aufsummieren $\sum_{j=1}^{n} a_{ij}b_{jk}$ ergibt das Matrixelement in der k-ten Spalte und i-ten Zeile der Produktmatrix."

**Beispiel 8.2.8.**

1. Obiges Schema an einem Zahlenbeispiel – mit den expliziten Zahlen für den ersten Spaltenvektor im Produkt – veranschaulicht:

$$\begin{array}{c|c}
 & \begin{pmatrix} 1 & -2 \\ 0 & -1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} \\ \hline
\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 2 & -2 \\ 2 & -1 & 0 & -3 \end{pmatrix} & \begin{pmatrix} 1 \cdot 1 + 2 \cdot 0 + (-1) \cdot 1 + 1 \cdot 1 = 1 & -5 \\ 0 \cdot 1 + 1 \cdot 0 + 2 \cdot 1 + (-2) \cdot 1 = 0 & 1 \\ 2 \cdot 1 + (-1) \cdot 0 + 0 \cdot 1 + (-3) \cdot 1 = -1 & -3 \end{pmatrix}
\end{array}$$

   oder

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 2 & -2 \\ 2 & -1 & 0 & -3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 \\ 0 & -1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -5 \\ 0 & 1 \\ -1 & -3 \end{pmatrix}$$

2. Matrixmultiplikation von Zeilen- und Spaltenvektoren:

$$(a_1, ..., a_n) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + ... + a_nb_n \in K$$

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \cdot (b_1, ..., b_n) = \begin{pmatrix} a_1b_1 & \cdots & a_1b_n \\ \vdots & & \vdots \\ a_nb_1 & \cdots & a_nb_n \end{pmatrix} \in M(n \times n, K).$$

3. Sind $A, B \in M(n \times n, K)$, so kann man $A \cdot B$ und $B \cdot A$ bilden, aber im allgemeinen ist $A \cdot B \neq B \cdot A$, wie zum Beispiel bei
$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

4. Die Einheitsmatrizen $E_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in M(n \times n; K)$ erfüllen $A \cdot E_n = A$ und $E_n \cdot B = B$ für alle $A \in M(m \times n, K)$ und $B \in M(n \times r, K)$.

5. Das Produkt einer Matrix $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} \in M(m \times n, K)$ mit einem Spaltenvektor $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ ist der Spaltenvektor $A \cdot \underline{x} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} \in K^m$.

**Korollar 8.2.9.** *$M_n(K) := M(n \times n, K)$ ist mit der Addition und Multiplikation von Matrizen ein Ring. In $M_n(K)$ ist $E_n$ das Einselement, dh. das neutrale Element von $(M_n(K), \cdot)$. Für $n > 1$ ist $M_n(K)$ nicht kommutativ.*

**Definition 8.2.10.** *Eine quadratische Matrix $A \in M_n(K)$ heißt* **invertierbar**, *falls es ein $A' \in M_n(K)$ gibt mit $A \cdot A' = A' \cdot A = E_n$. In diesem Fall schreibt man $A^{-1}$ statt $A'$ und nennt $A^{-1}$ die* **inverse Matrix** *von $A$. Eine nicht invertierbare Matrix heißt* **singulär**.

$$GL(n, K) := \{A \in M_n(K) : A \text{ invertierbar}\}. \tag{8.2.2}$$

**Bemerkung 8.2.11.** Ist $A \in M_n(K)$ invertierbar, so ist die Matrix $A'$ mit $A \cdot A' = A' \cdot A = E_n$ eindeutig bestimmt, denn für $A'' \in M_n(K)$ mit $A \cdot A'' = A'' \cdot A = E_n$ gilt $A'' = A'' \cdot (A \cdot A') = (A'' \cdot A) \cdot A' = E_n \cdot A' = A'$. Dies rechtfertigt die Schreibweise $A^{-1}$ in obiger Definition.

**Lemma 8.2.12.** *Sind $A, B \in GL(n, K)$, so gilt*
$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1} \tag{8.2.3}$$
$$(A^{-1})^{-1} = A. \tag{8.2.4}$$
*Damit ist $(GL(n, K), \cdot)$ eine Gruppe.*

*Beweis.* Zu $A, B \in GL(n, K)$ gibt es $A^{-1}, B^{-1} \in GL(n, K)$ und wegen
$$\begin{aligned}(B^{-1} \cdot A^{-1}) \cdot (A \cdot B) = B^{-1} \cdot (A^{-1} \cdot A) \cdot B = B^{-1} \cdot E_n \cdot B &= E_n \\ (A \cdot B) \cdot (B^{-1} \cdot A^{-1}) = A \cdot (B \cdot B^{-1}) \cdot A^{-1} = A \cdot E_n \cdot A^{-1} &= E_n\end{aligned}$$
ist $A \cdot B$ invertierbar und (8.2.3) gilt. Die Gleichung
$$A \cdot A^{-1} = A^{-1} \cdot A = E_n$$
läßt sich sowohl als Gleichung zur Bestimmung der inversen Matrix von $A$ als auch zur Bestimmung der inversen Matrix von $A^{-1}$ lesen, woraus (8.2.4) folgt. Da die Matrizenmultiplikation nach Lemma 8.2.6 assoziativ ist, zeigen (8.2.3) und (8.2.4), daß $(GL(n, K), \cdot)$ eine Gruppe ist. □

## 8.3 Die darstellende Matrix einer linearen Abbildung

**Beispiel 8.3.1.** Ist $A \in M(m\times n, K)$, so wird nach Lemma 8.2.6 durch $\begin{array}{rcl} F_A : K^n & \to & K^m \\ \underline{x} & \mapsto & A\underline{x} \end{array}$ eine lineare Abbildung definiert.

Als eine „Umkehrung“ dieses Beispiels, können wir jede lineare Abbildung zwischen endlichdimensionalen Vektorräumen als Matrix auffassen:

**Satz 8.3.2.** *Es sei $V$ ein $K$-Vektorraum mit Basis $\mathcal{A} = \{v_1, ..., v_n\}$ und $W$ ein $K$-Vektorraum mit Basis $\mathcal{B} = \{w_1, ..., w_m\}$. Dann gibt es zu jeder $K-$linearen Abbildung $f : V \to W$ genau eine Matrix $A = (a_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} \in M(m \times n, K)$, so daß*

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i \tag{8.3.1}$$

*für alle $j = 1, ..., n$ gilt. $M^{\mathcal{A}}_{\mathcal{B}}(f) := A$ heißt die* **darstellende Matrix von $f$ bezüglich der Basen $\mathcal{A}$ und $\mathcal{B}$**. *Die Abbildung*

$$\begin{array}{rcl} M^{\mathcal{A}}_{\mathcal{B}} : \mathrm{Hom}_K(V, W) & \to & M(m \times n, K) \\ f & \mapsto & M^{\mathcal{A}}_{\mathcal{B}}(f) \end{array} \tag{8.3.2}$$

*ist ein Isomorphismus von $K$-Vektorräumen.*

*Beweis.* Nach Satz 8.1.6 ist eine $K$-lineare Abbildung $f : V \to W$ bereits durch die Funktionswerte $f(v_1), ..., f(v_n)$ der Basisvektoren $v_1, ..., v_n$ eindeutig bestimmt. Da $\mathcal{B}$ eine Basis von $W$ ist, läßt sich jedes $f(v_j)$ in eindeutiger Weise in der Form (8.3.1) schreiben. Dies zeigt, daß es zu $f \in \mathrm{Hom}_K(V, W)$ genau ein $A = (a_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} =: M^{\mathcal{A}}_{\mathcal{B}}(f) \in M(m\times n, K)$ gibt, so daß (8.3.1) erfüllt ist. Damit ist $\begin{array}{rcl} M^{\mathcal{A}}_{\mathcal{B}} : \mathrm{Hom}_K(V, W) & \to & M(m \times n, K) \\ f & \mapsto & M^{\mathcal{A}}_{\mathcal{B}}(f) \end{array}$ bijektiv; Linearität von $M^{\mathcal{A}}_{\mathcal{B}}$ ist eine leichte Übung. $\square$

**Korollar 8.3.3.** *Ist $V = K^n$ und $W = K^m$ und sind $\mathcal{K}_n$ und $\mathcal{K}_m$ die Standardbasen von $K^n$ bzw. $K^m$, so ist*

$$\begin{array}{rcl} (M^{\mathcal{K}_n}_{\mathcal{K}_m})^{-1} : M(m \times n, K) & \to & \mathrm{Hom}_K(K^n, K^m) \\ A & \mapsto & \left( \begin{array}{rcl} F_A : K^n & \to & K^m \\ \underline{x} & \mapsto & A\underline{x} \end{array} \right) \end{array}$$

**Bemerkung 8.3.4.** Wenn nicht explizit etwas anderes gesagt wird – also andere Basen für $K^n$ bzw. $K^m$ angegeben werden – so identifiziert man via $(M^{\mathcal{K}_n}_{\mathcal{K}_m})^{-1}$ aus Korollar 8.3.3 eine Matrix mit einer linearen Abbildung. Man sollte sich aber schon an dieser Stelle im Klaren sein, daß in der Bezeichnung $M^{\mathcal{A}}_{\mathcal{B}}(f)$ die Angabe der beiden Basen $\mathcal{A}$ und $\mathcal{B}$ nicht nur schmückendes Beiwerk ist, sondern die darstellende Matrix von $f \in \mathrm{Hom}_K(V, W)$ ganz entscheidend von der Wahl von $\mathcal{A}$ und $\mathcal{B}$ abhängt. Ein spezielles „Basisdesign“ für eine möglichst einfache darstellende Matrix zeigt:

**Korollar 8.3.5.** *Ist $f : V \to W$ eine $K$-lineare Abbildung, $n = \dim_K V$, $m = \dim_K W$ und $r = \dim_K(f(V))$, dann gibt es eine Basis $\mathcal{A}$ von $V$ und eine Basis $\mathcal{B}$ von $W$, so daß*

$$M_{\mathcal{B}}^{\mathcal{A}}(f) = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right) \in M(m \times n, K) \tag{8.3.3}$$

*eine Blockmatrix bestehend aus der $r \times r$ Einheitsmatrix $E_r$ und Nullmatrizen in den restlichen Blöcken ist.*

*Beweis.* Zu jeder Basis $\{w_1, ..., w_r\}$ von $f(V)$ und $\{v_1, ..., v_s\}$ von $\mathrm{Kern} f$ erhält man nach Satz 8.1.7 durch Wahl von $u_1 \in f^{-1}(\{w_1\})$,..., $u_r \in f^{-1}(\{w_r\})$ eine Basis $\mathcal{A} = \{u_1, ..., u_r, v_1, ..., v_s\}$ von $V$. Ergänze $\{w_1, ..., w_r\}$ mit dem Basisergänzungssatz zu einer Basis $\mathcal{B} = \{w_1, ..., w_r, y_1, ..., y_{m-r}\}$ von $W$. Dann gilt offenbar (8.3.3). □

**Definition 8.3.6.** *Es sei $V$ ein $K$-Vektorraum mit Basis $\mathcal{B} = \{v_1, ..., v_n\}$ und $\{\underline{e_1}, ..., \underline{e_n}\}$ die Standardbasis von $K^n$. Den nach Satz 8.1.6 eindeutig bestimmten Isomorphismus*

$$\Phi_{\mathcal{B}} : K^n \to V \quad \textit{mit} \quad \Phi_{\mathcal{B}}(\underline{e_j}) = v_j \quad \textit{für } j = 1, ..., n$$

*nennt man das durch $\mathcal{B}$ definierte* **Koordinatensystem** *und zu $v \in V$ heißt* $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \Phi_{\mathcal{B}}^{-1}(v) \in K^n$ *der* **Koordinatenvektor** *zu $v \in V$.*

**Lemma 8.3.7.** *Sind $V, W$ $K$-Vektorräume, $\mathcal{A} = \{v_1, ..., v_n\}$ eine Basis von $V$ und $\mathcal{B} = \{w_1, ..., w_m\}$ eine Basis von $W$ und $f : V \to W$ $K-$linear, dann gilt*

$$F_{M_{\mathcal{B}}^{\mathcal{A}}(f)} = \Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}} \tag{8.3.4}$$

*Beweis.* Ist $\{\underline{e_1}, ..., \underline{e_n}\}$ die Standardbasis von $K^n$ und $(a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} = M_{\mathcal{B}}^{\mathcal{A}}(f)$, so folgt für $j = 1, ..., n$ aus den Definitionen

$$\Phi_{\mathcal{B}}(M_{\mathcal{B}}^{\mathcal{A}}(f)\underline{e_j}) = \Phi_{\mathcal{B}} \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \sum_{i=1}^{m} a_{ij} w_i$$

und

$$f(\Phi_{\mathcal{A}}(\underline{e_j})) = f(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Da $\Phi_{\mathcal{B}}$ ein Isomorphismus ist, folgt daraus $M_{\mathcal{B}}^{\mathcal{A}}(f)\underline{e_j} = (\Phi_{\mathcal{B}}^{-1} \circ f \circ \Phi_{\mathcal{A}})(\underline{e_j})$ für $j = 1, ..., n$, was nach Satz 8.1.6 reicht um (8.3.4) zu zeigen. □

**Bemerkung 8.3.8.** In der Situation von Lemma 8.3.7 enthält der $j$-te Spaltenvektor $M_{\mathcal{B}}^{\mathcal{A}}(f)\underline{e_j}$ der darstellenden Matrix $M_{\mathcal{B}}^{\mathcal{A}}(f)$ von $f$ bezüglich der Basen $\mathcal{A}$ und $\mathcal{B}$ den Koordinatenvektor des Bildes $f(v_j)$ geschrieben in der Basis $\mathcal{B}$.

**Beispiel 8.3.9.** $\mathcal{B} = \left( \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right) =: (\underline{v}_1, \underline{v}_2, \underline{v}_3)$ ist eine Basis [1] von $\mathbb{R}^3$.
Ist die lineare Abbildung $\begin{array}{rccc} F_A : & \mathbb{R}^3 & \to & \mathbb{R}^3 \\ & \underline{x} & \mapsto & A\underline{x} \end{array}$ bezüglich der Standardbasis gegeben durch die Matrix

$$A = \begin{pmatrix} 1 & 3 & -2 \\ -1 & 4 & -1 \\ 0 & -1 & 3 \end{pmatrix}$$

so bestimmt sich $M_{\mathcal{B}}^{\mathcal{B}}(F_A)$, indem man $F_A(\underline{v}_1)$, $F_A(\underline{v}_2)$, $F_A(\underline{v}_3)$ berechnet, in der Basis $\mathcal{B}$ schreibt und die Koeffizienten zu $M_{\mathcal{B}}^{\mathcal{B}}(F_A)$ anordnet:

$$\begin{pmatrix} 1 & 3 & -2 \\ -1 & 4 & -1 \\ 0 & -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 0 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + 0 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3 & -2 \\ -1 & 4 & -1 \\ 0 & -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -3 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 3 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + 0 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3 & -2 \\ -1 & 4 & -1 \\ 0 & -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ -1 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + 3 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

und daher ist

$$M_{\mathcal{B}}^{\mathcal{B}}(F_A) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

**Satz 8.3.10.** *Es seien $U, V, W$ endlichdimensionale $K$-Vektorräume mit Basen $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$, sowie $K$-lineare Abbildungen $g : U \to V$ und $f : V \to W$. Dann gilt*

$$M_{\mathcal{C}}^{\mathcal{A}}(f \circ g) = M_{\mathcal{C}}^{\mathcal{B}}(f) \cdot M_{\mathcal{B}}^{\mathcal{A}}(g), \tag{8.3.5}$$

*dh. die* **Hintereinanderausführung linearer Abbildungen entspricht dem Matrixprodukt ihrer darstellenden Matrizen**.

*Beweis.* Ist $u \in U$ und $\underline{x} := \Phi_{\mathcal{A}}^{-1}(u)$, dann gilt $\Phi_{\mathcal{B}}^{-1}(g(u)) = (\Phi_{\mathcal{B}}^{-1} \circ g \circ \Phi_{\mathcal{A}})(\underline{x}) = M_{\mathcal{B}}^{\mathcal{A}}(g)\underline{x}$ nach Lemma 8.3.7. Ebenso folgt

$$\begin{aligned} \Phi_{\mathcal{C}}^{-1}(f(g(u)) &= (\Phi_{\mathcal{C}}^{-1} \circ f \circ g \circ \Phi_{\mathcal{A}})(\underline{x}) = (\Phi_{\mathcal{C}}^{-1} \circ f \circ \Phi_{\mathcal{B}} \circ \Phi_{\mathcal{B}}^{-1} \circ g \circ \Phi_{\mathcal{A}})(\underline{x}) \\ &= (\Phi_{\mathcal{C}}^{-1} \circ f \circ \Phi_{\mathcal{B}})(M_{\mathcal{B}}^{\mathcal{A}}(g)\underline{x}) = M_{\mathcal{C}}^{\mathcal{B}}(f)(M_{\mathcal{B}}^{\mathcal{A}}(g)\underline{x}) \\ &= (M_{\mathcal{C}}^{\mathcal{B}}(f) \cdot M_{\mathcal{B}}^{\mathcal{A}}(g)) \cdot (\Phi_{\mathcal{A}}^{-1}(u)), \end{aligned}$$

ist also $n = \dim_K U$ und $m = \dim_K(W)$, so ist dadurch gezeigt:

$$M_{\mathcal{C}}^{\mathcal{A}}(f \circ g) = M_{\mathcal{K}_m}^{\mathcal{K}_n}(\Phi_{\mathcal{C}}^{-1} \circ (f \circ g) \circ \Phi_{\mathcal{A}}) = M_{\mathcal{C}}^{\mathcal{B}}(f) \cdot M_{\mathcal{B}}^{\mathcal{A}}(g). \qquad \square$$

[1]Da drei Vektoren im dreidimensionalen $\mathbb{R}-$Vektorraum $\mathbb{R}^3$ genau dann eine Basis bilden, wenn sie linear unabhängig sind, braucht man nur die lineare Unabhängigkeit zu zeigen – und das geht hier ganz schnell

**Satz 8.3.11.** *Es sei $n = \dim_K V$, $m = \dim_K W$ und $f : V \to W$ sei $K$-linear. Ferner seien $\mathcal{A}$ und $\mathcal{A}'$ Basen von $V$ und $\mathcal{B}$ und $\mathcal{B}'$ Basen von $W$. Dann gilt:*

$$M^{\mathcal{A}'}_{\mathcal{B}'}(f) = M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_W) M^{\mathcal{A}}_{\mathcal{B}}(f) \left(M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V)\right)^{-1}. \tag{8.3.6}$$

*Beweis.* Nach Lemma 8.3.7 ist in

$$\begin{tikzcd}
K^n \arrow[rrr, "F_{M^{\mathcal{A}}_{\mathcal{B}}(f)}"] \arrow[dr, "\Phi_{\mathcal{A}}"] \arrow[dd, "F_{M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V)}"'] & & & K^m \arrow[dl, "\Phi_{\mathcal{B}}"'] \arrow[dd, "F_{M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_W)}"] \\
& V \arrow[r, "f"] & W & \\
K^n \arrow[ur, "\Phi_{\mathcal{A}'}"'] \arrow[rrr, "F_{M^{\mathcal{A}'}_{\mathcal{B}'}(f)}"] & & & K^m \arrow[ul, "\Phi_{\mathcal{B}'}"]
\end{tikzcd}$$

jedes Dreieck und jedes Trapez für sich genommen ein **kommutatives Diagramm** [2] und damit ist das ganze Bild ein kommutatives Diagramm, insbesondere

$$F_{M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_W)} \circ F_{M^{\mathcal{A}}_{\mathcal{B}}(f)} = F_{M^{\mathcal{A}'}_{\mathcal{B}'}(f)} \circ F_{M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V)}.$$

Da $\mathrm{id}_V$ ein Isomorphismus ist, gilt dies nach (8.3.2) auch für $F_{M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V)}$ und damit folgt

$$F_{M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_W)} \circ F_{M^{\mathcal{A}}_{\mathcal{B}}(f)} \circ \left(F_{M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V)}\right)^{-1} = F_{M^{\mathcal{A}'}_{\mathcal{B}'}(f)}$$

bzw. (8.3.6) für die darstellenden Matrizen. □

## 8.4 Der Rang einer Matrix

**Definition 8.4.1.** *Es sei* $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \left(\underline{a_1}|\cdots|\underline{a_n}\right) = \begin{pmatrix} \vec{a_1} \\ \hline \vdots \\ \hline \vec{a_m} \end{pmatrix}$ *eine $m \times n$ Matrix über $K$, einmal mit den Spaltenvektoren $\underline{a_1}, ..., \underline{a_n} \in K^m$ und einmal mit den Zeilenvektoren $\vec{a_1}, ..., \vec{a_m} \in K^n$ geschrieben. Dann heißt*

$$A^T := \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix} \in M(n \times m, K) \tag{8.4.1}$$

[2] dh. jede mögliche Hintereinanderausführung von Funktionen in diesem Diagramm ergibt die gleiche Funktion, zB. $F \circ \Phi_{\mathcal{A}'} = \Phi_{\mathcal{B}'} \circ F_{M^{\mathcal{A}'}_{\mathcal{B}}(f)}$. Ein kommutatives Diagramm ist also nichts Neues, aber eine gute Möglichkeit, bei vielen Funktionen die Übersicht zu behalten

*die* **transponierte Matrix von** *A,*

$$\mathrm{Z}-\mathrm{Rang}(A) := \dim_K \mathrm{lin}(\{\vec{a_1}, ..., \vec{a_m}\}) \tag{8.4.2}$$

*der* **Zeilenrang** *von A,*

$$\mathrm{S}-\mathrm{Rang}(A) := \dim_K \mathrm{lin}(\{\underline{a_1}, ..., \underline{a_n}\}) \tag{8.4.3}$$

*der* **Spaltenrang** *von A und*

$$\mathrm{Rang}(A) := \dim_K(F_A(K^n)) \tag{8.4.4}$$

*der* **Rang** *von A.*

**Bemerkung 8.4.2.** Sind $A \in M(m\times n, K)$ und $B \in M(n\times p, K)$, so ist $(AB)^T = B^T A^T$, wie man aus den expliziten Formeln zur Matrixmultiplikation nachrechnet. Ferner ist $\mathrm{Z}-\mathrm{Rang}(A) = \mathrm{S}-\mathrm{Rang}(A^T)$.

**Satz 8.4.3.** *Für $A \in M(m \times n, K)$ gilt*

$$\mathrm{Z}-\mathrm{Rang}(A) = \mathrm{S}-\mathrm{Rang}(A) = \mathrm{Rang}(A). \tag{8.4.5}$$

*Beweis.* Für $A = (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} = (\underline{a_1}|...|\underline{a_n}) \in M(m \times n, K)$ und $\begin{array}{rcl} F_A : K^n & \to & K^m \\ \underline{x} & \mapsto & A\underline{x} \end{array}$ gilt $F_A(\underline{e_j}) = \underline{a_j}$ für die Basisvektoren $\underline{e_1}, ..., \underline{e_n}$ der Standardbasis von $K^n$. Damit ist $\mathrm{S}-\mathrm{Rang}(A) = \mathrm{Rang}(A)$. Nach Korollar 8.3.5 gibt es eine Basis $\mathcal{A}$ von $K^n$ und eine Basis $\mathcal{B}$ von $K^m$, so daß mit $r := \mathrm{Rang}(A)$

$$B := M_{\mathcal{B}}^{\mathcal{A}}(F_A) = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right) \in M(m \times n, K)$$

ist. Dann sind die ersten $r$ Zeilenvektoren (bzw. Spaltenvektoren) von $B$ linear unabhängig und deshalb gilt $\mathrm{Z}-\mathrm{Rang}(B) = r = \mathrm{S}-\mathrm{Rang}(B) = \mathrm{Rang}(B)$. Nach Satz 8.3.11 gibt es $S \in GL(m, K)$ und $T \in GL(n, K)$ mit $B = SAT^{-1}$. Wegen $T^{-1} \in GL(n.K)$ gilt $\mathrm{Kern}(F_{T^{-1}}) = \{\underline{0}\}$ für $\begin{array}{rcl} F_{T^{-1}} : K^n & \to & K^n \\ \underline{y} & \mapsto & T^{-1}\underline{y} \end{array}$ ; nach der Dimensionsformel folgt

$$n = \dim_K(F_{T^{-1}}(K^n)) + \dim_K(\mathrm{Kern}(F_{T^{-1}})) = \dim_K(F_{T^{-1}}(K^n)),$$

also ist $F_{T^{-1}}(K^n) = K^n$ und damit

$$\mathrm{Rang}(AT^{-1}) = \dim_K(F_{AT^{-1}}(K^n)) = \dim_K((F_A \circ F_{T^{-1}})(K^n)) = \dim_K(F_A(K^n)) = \mathrm{Rang}(A). \tag{8.4.6}$$

$V := F_A(K^n) \subseteq K^m$ ist ein Untervektorraum, also ist die Einschränkung $\begin{array}{rcl} F_S|_V : V & \to & K^m \\ \underline{y} & \mapsto & S\underline{y} \end{array}$ von $F_S$ auf $V$ wieder linear und wegen $\mathrm{Kern}(F_S|_V) \subseteq \mathrm{Kern}(F_S) = \{\underline{0}\}$ folgt aus der Dimensionsformel:

$$\begin{aligned} \mathrm{Rang}(A) &= \dim_K(V) = \dim_K((F_S|_V)(V)) + \dim_K(\mathrm{Kern}(F_S|_V)) = \dim_K((F_S|_V)(V)) \\ &= \dim_K((F_S \circ F_A)(K^n)) = \mathrm{Rang}(SA). \end{aligned} \tag{8.4.7}$$

Aus (8.4.6) und (8.4.7) folgt $\mathrm{S}-\mathrm{Rang}(B) = \mathrm{Rang}(SAT^{-1}) = \mathrm{Rang}(A) = \mathrm{S}-\mathrm{Rang}(A)$. Wegen $\mathrm{Z}-\mathrm{Rang}(A) = \mathrm{S}-\mathrm{Rang}(A^T) = \mathrm{S}-\mathrm{Rang}(B^T) = \mathrm{Z}-\mathrm{Rang}(B)$ folgt (8.4.5). □

**Satz 8.4.4.** *Es sei $A \in M_n(K)$, dann sind äquivalent:*

*a)* $A \in GL(n, K)$

*b)* $\begin{array}{rcl} F_A : K^n & \to & K^n \\ \underline{x} & \mapsto & A\underline{x} \end{array}$ *ist bijektiv.*

*c)* $\mathrm{Rang}(A) = n$

*d)* $\mathrm{Kern}(F_A) = \{\underline{0}\}$

*e) Für jedes $\underline{b} \in K^n$ hat $A\underline{x} = \underline{b}$ genau eine Lösung.*

*Beweis.* Offenbar sind b) und e) äquivalent.

Ist $A \in GL(n, K)$, so ist $F_A \circ F_{A^{-1}} = F_{A^{-1}} \circ F_A = \mathrm{id}_{K^n}$, also $F_A$ bijektiv. Ist umgekehrt $F_A$ bijektiv, so existiert die Umkehrfunktion $F_A^{-1}$, die nach Satz 8.1.3 $K$-linear ist. In der Standardbasis $\mathcal{K}_n$ des $K^n$ folgt dann nach Satz 8.3.10 für die darstellenden Matrizen:

$$A \cdot M_{\mathcal{K}_n}^{\mathcal{K}_n}(F_A^{-1}) = M_{\mathcal{K}_n}^{\mathcal{K}_n}(F_A) \cdot M_{\mathcal{K}_n}^{\mathcal{K}_n}(F_A^{-1}) = M_{\mathcal{K}_n}^{\mathcal{K}_n}(F_A \circ F_A^{-1}) = E_n = ... = M_{\mathcal{K}_n}^{\mathcal{K}_n}(F_A^{-1}) \cdot A,$$

was $a) \Leftrightarrow b)$ zeigt.

Da $F_A$ genau dann surjektiv ist, wenn $\mathrm{Rang}(A) = n$ und $F_A$ genau dann injektiv ist, wenn $\mathrm{Kern}(F_A) = \{\underline{0}\}$ ist, folgt aus der Dimensionsformel $n = \dim_K \mathrm{Kern}(F_A) + \mathrm{Rang}(A)$ die Äquivalenz von $F_A$ surjektiv mit $F_A$ injektiv. □

**Definition 8.4.5.** *$A, B \in M(m \times n, K)$ heißen* **äquivalent***, wenn es $S \in GL(m, K)$ und $T \in GL(n, K)$ gibt mit $B = SAT^{-1}$. $A, B \in M_n(K)$ heißen* **ähnlich***, wenn es $S \in GL(n, K)$ gibt mit $B = SAS^{-1}$.*

**Bemerkung 8.4.6.** Aus Satz 8.3.11 folgt, daß $A, B \in M(m \times n, K)$ wenn sie dieselbe lineare Abbildung in verschiedenen Basen beschreiben, äquivalent sind. Ebenso sind $A, B \in M_n(K)$ ähnlich, wenn sie in verschiedenen Basen denselben Endomorphismus beschreiben, dh. wenn es zwei Basen $\mathcal{A}$ und $\mathcal{B}$ von $V$ und einen Endomorphismus $F : V \to V$ mit $A = M_{\mathcal{A}}^{\mathcal{A}}(F)$ und $B = M_{\mathcal{B}}^{\mathcal{B}}(F)$ gibt.

**Korollar 8.4.7.** *$A, B \in M(m \times n, K)$ sind genau dann äquivalent, wenn $\mathrm{Rang}(A) = \mathrm{Rang}(B)$; insbesondere sind für $\mathrm{Rang}(A) = \mathrm{Rang}(B) = r$ die Matrizen $A$ und $B$ zur Blockmatrix $\left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right)$ äquivalent.*

*Beweis.* $\Rightarrow$ wurde beim Beweis von Satz 8.4.3 gezeigt. Nach Satz 8.3.11 und Korollar 8.3.5 sind $A$ und $B$ zu $\left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right)$ äquivalent, was auch die umgekehrte Richtung zeigt. □

# Kapitel 9

# Lineare Gleichungssysteme

## 9.1 Affine Unterräume

**Definition 9.1.1.** *Es sei $V$ ein $K$-Vektorraum, dann heißt $X \subseteq V$ ein* **affiner Unterraum**, *wenn $X = \emptyset$ oder wenn es ein $v \in V$ und einen Untervektorraum $W \subseteq V$ gibt mit*

$$X = v + W = \{v + w : w \in W\}.$$

**Lemma 9.1.2.** *Ist $\emptyset \neq X = v + W \subseteq V$ ein affiner Unterraum, so gilt:*

*a) Für jedes $x \in X$ gilt $X = x + W$.*

*b) Ist $x \in X$ und $\tilde{W} \subseteq V$ ein Untervektorraum mit $v + W = x + \tilde{W}$, so gilt $W = \tilde{W}$ und $x - v \in W$.*

*Beweis.*

a) Ist $x = v + u \in X$ mit $u \in W$ und $y = v + \tilde{w} \in X$ mit $\tilde{w} \in W$, dann ist $y = (x-u)+\tilde{w} = x+(\tilde{w}-u) \in x+W$, dh. $X = v+W \subseteq x+W$. Zu $y \in x+W$ gibt es $\tilde{w} \in W$ mit $y = x+\tilde{w} = (v+u)+\tilde{w} = v+(u+\tilde{w}) \in v+W$, dh. $x+W \subseteq X = v+W$.

b) Betrachte $(v + W) - (x + \tilde{W}) := \{v + w - (x + \tilde{w}) : w \in W, \tilde{w} \in \tilde{W}\}$. Wegen $v + W = x + \tilde{W}$ gibt es für jedes $x + \tilde{w} \in x + \tilde{W}$ ein $w' \in W$ mit $x + \tilde{w} = v + w'$, also ist

$$\begin{aligned}(v + W) - (x + \tilde{W}) &= \{v + w - (x + \tilde{w}) : w \in W, \tilde{w} \in \tilde{W}\} \\ &= \{v + w - (v + w') : w, w' \in W\} = W\end{aligned}$$

und analog zeigt man $(v + W) - (x + \tilde{W}) = \tilde{W}$, daher $W = \tilde{W}$. Damit ist $v + W = x + W$, also $x = x + 0 = v + u \in v + W$, dh. $x - v = u \in W$.

□

**Definition 9.1.3.** *Ist $\emptyset \neq X = v + W$ ein affiner Unterraum von $V$, so heißt* $\dim_K X := \dim_K W$ *die* **Dimension** *von $X$.*

## 9.2 Lösen linearer Gleichungssyteme

**Definition 9.2.1** (Homogene und inhomogene Gleichungssysteme)**.** *Sind $a_{ij} \in K$ für $i = 1, ..., m$ und $j = 1, ..., n$ vorgegeben, so nennt man das Gleichungssystem*

$$\begin{array}{rcl} a_{11}x_1 + ... + a_{1n}x_n & = & 0 \\ a_{21}x_1 + ... + a_{2n}x_n & = & 0 \\ \vdots \qquad\qquad \vdots & & \vdots \\ a_{m1}x_1 + ... + a_{mn}x_n & = & 0 \end{array} \tag{9.2.1}$$

*ein* **homogenes lineares Gleichungssystem** *(in $n$ Unbekannten $x_1, ..., x_n$ bestehend aus $m$ Gleichungen). Die Matrix $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ heißt die* **Koeffizientenmatrix** *von (9.2.1). Setzt man $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, so schreibt sich das Gleichungssystem (9.2.1) als*

$$A\underline{x} = \underline{0}$$

*Ist noch $\underline{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m \backslash \{\underline{0}\}$ gegeben, so nennt man*

$$A\underline{x} = \underline{b} \tag{9.2.2}$$

*oder ausgeschrieben*

$$\begin{array}{rcl} a_{11}x_1 + ... + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + ... + a_{2n}x_n & = & b_2 \\ \vdots \qquad\qquad \vdots & & \vdots \\ a_{m1}x_1 + ... + a_{mn}x_n & = & b_m \end{array}$$

*ein* **inhomogenes lineares Gleichungssystem***. Zu $\underline{b} \in K^m$ heißt*

$$L\ddot{o}s(A, \underline{b}) := \{\underline{x} \in K^n : A\underline{x} = \underline{b}\} \tag{9.2.3}$$

*der* **Lösungsraum** *von $A\underline{x} = \underline{b}$.*

**Lemma 9.2.2.** *Es sei $A \in M(m \times n, K)$. Der Lösungsraum $L\ddot{o}s(A, \underline{0})$ des homogenen linearen Gleichungssystems $A\underline{x} = \underline{0}$ ist ein Untervektorraum. Für $\underline{b} \neq \underline{0}$ ist der Lösungsraum $L\ddot{o}s(A, \underline{b})$ des inhomogenen linearen Gleichungssystems $A\underline{x} = \underline{b}$ ein affiner Unterraum.*

*Beweis.* Betrachte die lineare Abbildung $\begin{array}{rcl} F_A : K^n & \to & K^m \\ \underline{x} & \mapsto & A\underline{x} \end{array}$, dann ist $\text{Lös}(A, \underline{0}) = \text{Kern} F_A$ ein Untervektorraum. Es sei nun $\underline{b} \neq \underline{0}$ und $\text{Lös}(A, \underline{b}) \neq \emptyset$. Sind dann $\underline{x}, \underline{v} \in K^n$ mit

$A\underline{v} = \underline{b}$ und $A\underline{x} = \underline{0}$, so folgt $A(\underline{v} + \underline{x}) = A\underline{v} + A\underline{x} = \underline{b} + \underline{0} = \underline{b}$, was die Inklusion $\underline{v} + \text{Lös}(A, \underline{0}) \subseteq \text{Lös}(A, \underline{b})$ zeigt. Wählt man andererseits ein $\underline{u} \in \text{Lös}(A, \underline{b})$, also $A\underline{u} = \underline{b}$, so können wir $\underline{u} = \underline{v} + (\underline{u} - \underline{v})$ schreiben und erhalten $A(\underline{u} - \underline{v}) = A\underline{u} - A\underline{v} = \underline{b} - \underline{b} = \underline{0}$, das heißt $\underline{u} - \underline{v} \in \text{Lös}(A, \underline{0})$, was die umgekehrte Inklusion $\text{Lös}(A, \underline{b}) \subseteq \underline{v} + \text{Lös}(A, \underline{0})$ zeigt. □

**Bemerkung 9.2.3.** $A\underline{x} = \underline{0}$ hat immer (mindestens) eine Lösung, nämlich $\underline{x} = \underline{0}$. Es gibt Beispiele für inhomogene lineare Gleichungssysteme $A\underline{x} = \underline{b}$, die keine Lösung besitzen, etwa $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ und $\underline{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Auskunft darüber, wann ein inhomogenes lineares Gleichungssystem Lösungen besitzt, gibt uns:

**Satz 9.2.4.** *Es sei $A \in M(m \times n, K)$ und $\underline{b} \in K^m$. Gilt für die* **erweiterte Koeffizientenmatrix** $(A|\underline{b}) \in M(m \times (n+1), K)$:

- $\text{Rang}(A|\underline{b}) > \text{Rang}(A)$, *dann ist Lös*$(A, \underline{b}) = \emptyset$.
- $\text{Rang}(A|\underline{b}) = \text{Rang}(A)$, *dann ist Lös*$(A, \underline{b})$ *ein* $n - \text{Rang}(A)$ *dimensionaler affiner Unterraum von $K^n$ und für jedes $\underline{v} \in$ Lös$(A, \underline{b})$ gilt:*

$$\textit{Lös}(A, \underline{b}) = \underline{v} + \textit{Lös}(A, \underline{0}). \tag{9.2.4}$$

*Beweis.* Ist $A = (\underline{a_1}| \cdots |\underline{a_n})$, so ist

$$\begin{aligned} \text{Rang}(A|\underline{b}) &= \text{S} - \text{Rang}(A|\underline{b}) = \dim_K \text{lin}(\{\underline{a_1}, ..., \underline{a_n}, \underline{b}\}) \geq \dim_K \text{lin}(\{\underline{a_1}, ..., \underline{a_n}\}) \\ &= \text{S} - \text{Rang}(A) = \text{Rang}(A), \end{aligned}$$

daher kann nur eine der beiden obigen Möglichkeiten eintreten.

- Ist nun $\text{Rang}(A|\underline{b}) > \text{Rang}(A)$, dann ist $\underline{b} \notin \text{lin}(\{\underline{a_1}, ..., \underline{a_n}\})$, dh. für jedes $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ ist $\underline{b} \neq \underline{a_1}x_1 + ... + \underline{a_n}x_n = A\underline{x}$, also ist $\text{Lös}(A|\underline{b}) = \emptyset$.
- Ist $\text{Rang}(A|\underline{b}) = \text{Rang}(A)$, so ist $\underline{b} \in \text{lin}(\{\underline{a_1}, ..., \underline{a_n}\})$, dh. es gibt ein $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ mit $\underline{b} = \underline{a_1}x_1 + ... + \underline{a_n}x_n = A\underline{x}$, also ist $\text{Lös}(A|\underline{b}) \neq \emptyset$. (9.2.4) haben wir im Fall $\text{Lös}(A, \underline{b}) \neq \emptyset$ bereits im Beweis von Lemma 9.2.2 gezeigt. Da $\text{Lös}(A, \underline{0})$ ein $n - \text{Rang}(A)$ dimensionaler Untervektorraum von $K^n$ ist folgt die Behauptung über die Dimension von $\text{Lös}(A, \underline{b})$. □

**Definition 9.2.5.** $B = (b_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in M(m \times n, K)$ *heißt in* **Zeilenstufenform**, *wenn es*

- *ein $r \in \{0, 1, ..., m\}$ gibt, so daß in den Zeilen von $B$ mit Indizes $1$ bis $r$ von* **0** *verschiedene Zeilenvektoren stehen und in den Zeilen mit Indizes $r+1$ bis $m$ jeweils der Nullvektor steht.*
- *Für $i \in \{1, ..., r\}$ sei $j_i := \min\{j \in \{1, ..., n\} : b_{ij} \neq 0\}$, dann gilt die Stufenbedingung $j_1 < j_2 < .... < j_r$.*

*In diesem Fall heißen die Elemente* $b_{1j_1} \neq 0, ..., b_{rj_r} \neq 0$ **Pivotelemente** *von* $B$.

**Rezept 9.2.6.** Ist bei einem homogenen linearen Gleichungssystem $B\underline{x} = \underline{0}$ die Koeffizientenmatrix $B$ in Zeilenstufenform, also

$$B = \begin{pmatrix} 0 & \cdots & 0 & b_{1j_1} & \cdots & & & & & & \\ 0 & & & & 0 & b_{2j_2} & \cdots & & & & \\ 0 & & & & & & 0 & b_{3j_3} & \cdots & & \\ 0 & & & & & & & & & \ddots & \\ 0 & & & & & & & & 0 & b_{rj_r} & \cdots \\ 0 & & & & & & & & & & 0 \\ \vdots & & & & & & & & & & \vdots \\ 0 & & & & & & & & & & 0 \end{pmatrix}$$

mit den Pivotelementen $b_{1,j_1} \neq 0, ..., b_{r,j_r} \neq 0$ so sieht man, welche Unbekannten bestimmt werden können und welche mit einem freien Parameter in die Lösung eingehen: Alle Unbekannten $x_{j_1}, ..., x_{j_r}$, deren Koeffizienten $b_{1,j_1}, ..., b_{r,j_r}$ Pivotelemente sind, können wir bestimmen: Beginnend mit der Gleichung in der $r$-ten Zeile von $B$, also

$$b_{r,j_r}x_{j_r} + b_{r,j_r+1}x_{j_r+1} + ... + b_{r,n}x_n = 0$$

lösen wir nach $x_{j_r}$ auf, setzen das Ergebnis in die Gleichung der $(r-1)$-ten Zeile ein und lösen dann nach $x_{j_{r-1}}$ auf, etc... . Alle anderen Unbekannten gehen als freie Parameter in die Lösung ein. Das ist keine Überraschung, denn offenbar ist $\mathrm{Z-Rang}(B) = r$, also nach der Dimensionsformel und Satz 8.4.3:

$$\begin{aligned} \dim_K \mathrm{Lös}(B, \underline{0}) &= \dim_K \mathrm{Kern}(F_B) = n - \dim_K(F(K^n)) = n - \mathrm{Rang}(B) \\ &= n - (\mathrm{Z-Rang}(B)) = n - r \end{aligned}$$

**Beispiel 9.2.7.** $B = \begin{pmatrix} 1 & 2 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 2 & 4 \end{pmatrix}$ ist in Zeilenstufenform, so erhalten wir beim Lösen von $B\underline{x} = \underline{0}$ oder ausgeschrieben

$$\begin{array}{ccccccc} x_1 & +2x_2 & +3x_3 & & & = & 0 \\ & & +x_3 & & +5x_5 & = & 0 \\ & & & 2x_4 & +4x_5 & = & 0 \end{array}$$

die aufgelösten Gleichungen $\begin{array}{rcl} x_4 & = & -2x_5 \\ x_3 & = & -5x_5 \\ x_1 & = & -2x_2 + 15x_5 \end{array}$ und damit

$$\mathrm{Lös}(B, \underline{0}) = \{(-2\lambda + 15\mu, \lambda, -5\mu, -2\mu, \mu) : \lambda, \mu \in K\}$$

**Definition 9.2.8.** *Es sei* $A = \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{pmatrix} \in M(m \times n, K)$*, dann heißen*

- *Multiplikation des i-ten Zeilenvektors $\vec{a_i}$ mit $\lambda \in K\backslash\{0\}$*
- *Addition des $\lambda$-fachen des i-ten Zeilenvektors zum l-ten Zeilenvektor ($l \neq i$)*
- *Vertauschen des i-ten Zeilenvektors $\vec{a_i}$ mit dem l-ten Zeilenvektor $\vec{a_l}$*

**elementare Zeilenumformungen**.

**Bemerkung 9.2.9.** Elementare Zeilenumformungen können durch Linksmultiplikation mit invertierbaren Matrizen geschrieben werden:

- Multiplikation der i-ten Zeile mit $\lambda \in K\backslash\{0\}$:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,n} \\ \lambda a_{i1} & \cdots & \lambda a_{in} \\ a_{i+1,1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = S_i(\lambda) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

$$\text{mit } S_i(\lambda) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} (\leftarrow \text{i-te Zeile})$$

- Addition des $\lambda$-fachen der j-ten Zeile zur i-ten Zeile:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} + \lambda a_{j1} & \cdots & a_{in} + \lambda a_{jn} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = Q_i^j(\lambda) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

$$\text{mit } Q_i^j(\lambda) = \begin{array}{c} \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & 1 & \cdots & \lambda & & \\ & & & & \ddots & \vdots & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix} \\ j\text{-te Spalte} \end{array} \quad i-teZeile$$

- Vertauschen der i-ten Zeile mit der j-ten Zeile:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = P_i^j \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

$$\text{mit } P_i^j = \begin{array}{c} \begin{pmatrix} 1 & & & | & & & & | & & & \\ & \ddots & & | & & & & | & & & \\ & & 1 & | & & & & | & & & \\ - & - & - & 0 & - & - & - & 1 & - & - & - \\ & & & | & 1 & & & | & & & \\ & & & | & & \ddots & & | & & & \\ & & & | & & & 1 & | & & & \\ - & - & - & 1 & - & - & - & 0 & - & - & - \\ & & & | & & & & | & 1 & & \\ & & & | & & & & | & & \ddots & \\ & & & | & & & & | & & & 1 \end{pmatrix} \\ \text{i-te Spalte} \qquad \text{j-te Spalte} \end{array} = (P_i^j)^{-1}$$

Für die diese **Elementarmatrizen** gilt offenbar

$$(S_i(\lambda))^{-1} = S_i\left(\frac{1}{\lambda}\right), \quad (Q_i^j(\lambda))^{-1} = Q_i^j(-\lambda), \quad (P_i^j)^{-1} = P_i^j. \tag{9.2.5}$$

**Satz 9.2.10.** *Ist $A \in M(m \times n, K)$, so läßt sich $A$ durch das Vertauschen von Zeilen und wiederholtes Addieren eines Vielfachen einer Zeile zu einer anderen Zeile in eine Matrix $B$ in Zeilenstufenform überführen.*

*Beweis.* durch Induktion nach der Zahl $n$ der Spaltenvektoren.

Induktionsanfang $n = 1$: Es sei $A = (\underline{a_1}) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} \in M(m \times 1, K)$. Ist $A \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, so kann man durch Zeilenvertauschung zu $A' = \begin{pmatrix} a'_{11} \\ \vdots \\ a'_{m1} \end{pmatrix}$ mit $a'_{11} \neq 0$ kommen und dann lassen sich alle anderen Einträge $a'_{i1}$, $i = 2, ..., m$ durch Addition von $-\frac{a'_{i1}}{a'_{11}}$ mal der ersten Zeile von $A$ eliminieren, dh. $B = \begin{pmatrix} a'_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ ist die Zeilenstufenform von $A$.

Induktionsschluß $n \to n+1$: Es sei $A = (\underline{a_1}|\underline{a_2}|\cdots|\underline{a_{n+1}}) \in M(m \times (n+1), K)$. Ist $\underline{a_1} = \underline{0}$, so läßt sich $\tilde{A} = (\underline{a_2}|\cdots|\underline{a_{n+1}}) \in M(m \times n, K)$ nach Induktionsvoraussetzung durch elementare Zeilenumformungen in Zeilenstufenform überführen, das ändert nichts am ersten Spaltenvektor und daher läßt sich auch $A$ durch elementare Zeilenumformungen in Zeilenstufenform bringen. Ist $\underline{a_1} \neq \underline{0}$, so läßt sich $A$ wie beim Induktionsanfang $n = 1$ auf die Form $A' = \left(\begin{array}{c|ccc} a'_{11} & a'_{12} & \cdots & a'_{1(n+1)} \\ \hline 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{array}\right)$ bringen. Nach Induktionsvoraussetzung läßt sich $\tilde{A} \in M((m-1) \times n, K)$ durch elementare Zeilenumformungen auf Zeilenstufenform bringen; diese Umformungen ändern die erste Spalte nicht, daher läßt sich auch $A$ durch elementare Zeilenumformungen auf Zeilenstufenform bringen. □

**Satz 9.2.11.** *Sei $A \in M(m \times n, K)$, $\underline{b} \in K^m$ und $S \in GL(m, K)$, dann gilt:*

$$Lös(A, \underline{b}) = Lös(SA, S\underline{b}). \tag{9.2.6}$$

*Beweis.* Wegen $S \in GL(m, K)$ ist die lineare Abbildung $F_S : K^m \to K^m$ ein Isomorphismus; also $\text{Rang}(A|\underline{b}) = \text{Rang}(SA|S\underline{b})$ und $\text{Rang}(A) = \text{Rang}(SA)$ nach Lemma 8.1.5. Aus Satz 9.2.4 folgt die Äquivalenz von $\emptyset = \text{Lös}(A, \underline{b})$ und $\emptyset = \text{Lös}(SA, S\underline{b})$.

Es sei also $\underline{u} \in \text{Lös}(A, \underline{b}) \neq \emptyset$, also $A\underline{u} = \underline{b}$, dann ist auch $SA\underline{u} = S\underline{b}$, dh. $\text{Lös}(A, \underline{b}) \subseteq \text{Lös}(SA, S\underline{b})$. Ist $\underline{v} \in \text{Lös}(SA, S\underline{b})$, also $SA\underline{v} = S\underline{b}$, dann ist $S^{-1}SA\underline{v} = A\underline{v} = S^{-1}S\underline{b} = \underline{b}$, dh. $\text{Lös}(SA, S\underline{b}) \subseteq \text{Lös}(A, \underline{b})$. □

**Rezept 9.2.12.** Es sei $A \in M(m \times n, K)$ und $\underline{b} \in K^m$. Zum Lösen von $A\underline{x} = \underline{b}$

- Bringe die erweiterte Koeffizientenmatrix $(A|\underline{b}) \in M(m \times (n+1), K)$ durch elementare Zeilenumformungen auf Zeilenstufenform.
- Ist $(SA|S\underline{b}) = \left(\begin{array}{cccccccccc|c} 0 & \cdots & 0 & c_{1j_1} & \cdots & & & & & & d_1 \\ 0 & & & & 0 & c_{2j_2} & \cdots & & & & d_2 \\ 0 & & & & & & 0 & c_{3j_3} & \cdots & & d_3 \\ 0 & & & & & & & & \ddots & & \vdots \\ 0 & & & & & & & 0 & c_{rj_r} & \cdots & d_r \\ 0 & & & & & & & & & 0 & d_{r+1} \\ 0 & & & & & & & & & 0 & 0 \\ \vdots & & & & & & & & & \vdots & \vdots \\ 0 & & & & & & & & & 0 & 0 \end{array}\right)$ eine Zeilenstufenform von $(A|b)$ mit den Pivotelementen $c_{1j_1}, \ldots, c_{rj_r} \neq 0$, so ist offenbar $r = \text{Rang}(A)$ und es gibt zwei Fälle:

  a) $d_{r+1} \neq 0$, dann ist $\text{Rang}(SA) = \text{Rang}(A) = r < \text{Rang}(SA|S\underline{b}) = r+1$, also $\text{Lös}(A|\underline{b}) = \text{Lös}(SA|S\underline{b}) = \emptyset$.

b) $d_{r+1} = ... = d_m = 0$, dann ist $\text{Lös}(A|\underline{b})$ ein $(n-r)$-dimensionaler affiner Unterraum. Eine Lösung von $A\underline{x} = \underline{b}$ erhält man, indem man die $r \times r$-Matrix

$$C := \begin{pmatrix} c_{1j_1} & c_{1j_2} & \cdots & c_{1j_r} \\ & c_{2j_2} & & \vdots \\ & & \ddots & \\ & & & c_{rj_r} \end{pmatrix}$$

betrachtet und für $\underline{y} := \begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_r} \end{pmatrix}$ und $\underline{d} := \begin{pmatrix} d_1 \\ \vdots \\ d_r \end{pmatrix}$ das Gleichungssystem $C\underline{y} = \underline{d}$ löst, indem man von unten her einsetzt:

$$\begin{aligned} x_{j_r} &= \frac{d_r}{c_{rj_r}} \\ x_{j_{r-1}} &= \frac{1}{c_{r-1,j_{r-1}}}(d_{r-1} - c_{r-1,j_r}x_{j_r}) = \frac{1}{c_{r-1,j_{r-1}}}(d_{r-1} - \frac{c_{r-1,j_r}d_r}{c_{rj_r}}) \\ &\vdots \end{aligned}$$

Dann ist

$$\underline{x} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x_{j_1} \\ 0 \\ \vdots \\ 0 \\ x_{j_2} \\ 0 \\ \vdots \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow j_1\text{-te Zeile} \\ \\ \\ \\ \leftarrow j_2\text{-te Zeile} \\ \\ \\ \end{matrix}$$

eine Lösung von $A\underline{x} = \underline{b}$. Den Lösungsraum $\text{Lös}(A,\underline{0}) = \text{Lös}(SA,\underline{0})$ von $SA\underline{x} = \underline{0}$ bestimmt man wie in Rezept 9.2.6, denn $SA$ ist eine Matrix in Zeilenstufenform und dann ist

$$\text{Lös}(A,\underline{b}) = \underline{x} + \text{Lös}(A,\underline{0}).$$

**Beispiel 9.2.13.**

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 1 \\ x_1 - 2x_2 - x_4 &= 2 \\ x_3 + x_4 &= -1 \end{aligned}$$

Wir formen die erweiterte Koeffizientenmatrix dieses Systems

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 1 \\ 1 & -2 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & -1 \end{array}\right)$$

auf Zeilenstufenform um. Dazu subtrahieren wir zuerst die 1. Zeile von der 2. Zeile und erhalten: $\left(\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 \\ 0 & 0 & 1 & 1 & -1 \end{array}\right)$ und addieren dann die 2. Zeile zur 3. Zeile und erhalten so eine Zeilenstufenform

$$(SA|S\underline{b}) = \left(\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Wir sind also bei Fall b) in Rezept 9.2.12 angelangt, also hat $A\underline{x} = \underline{b}$ eine Lösung. Nur die Unbekannten $x_1$ und $x_3$ haben in einer der Gleichungen ein Pivotelement als Koeffizienten; das Gleichungssystem $C\underline{y} = \underline{d}$ aus Rezept 9.2.12 lautet also hier

$$\begin{aligned} x_1 + x_3 &= 1 \\ -x_3 &= 1 \end{aligned}$$

was die Lösung $x_3 = -1$ und $x_1 = 1 - x_3 = 2$ hat und damit ist $\underline{x} = \begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \end{pmatrix} \in \text{Lös}(A, \underline{b})$.

Der Lösungsraum $\text{Lös}(A, \underline{0})$ des homogenen Gleichungssystems $A\underline{x} = \underline{0}$ folgt ebenso aus obiger Zeilenstufenform als

$$\text{Lös}(A, \underline{0}) = \{\begin{pmatrix} 2\lambda_1 + \lambda_2 \\ \lambda_1 \\ -\lambda_2 \\ \lambda_2 \end{pmatrix} \in K^4 : \lambda_1, \lambda_2 \in K\}$$

und damit sind die Lösungen von $A\underline{x} = \underline{b}$ gerade

$$\text{Lös}(A, \underline{b}) = \underline{x} + \text{Lös}(A, \underline{0}) = \{\begin{pmatrix} 2 + 2\lambda_1 + \lambda_2 \\ \lambda_1 \\ -1 - \lambda_2 \\ \lambda_2 \end{pmatrix} \in K^4 : \lambda_1, \lambda_2 \in K\}.$$

## 9.3 Invertieren von Matrizen

**Satz 9.3.1.** *Jede invertierbare Matrix $A \in GL(n, K)$ ist ein Produkt aus Elementarmatrizen (dh. von Matrizen der Form $S_i(\lambda), Q_i^j(\lambda)$ oder $P_i^j$ aus Bemerkung 9.2.9 mit $\lambda \in K\backslash\{0\}$).*

*Beweis.* Durch elementare Zeilenumformungen können wir $A$ nach Satz 9.2.10 auf Zeilenstufenform $B = SA$ mit $S \in GL(n, K)$ bringen. Weil $A$ invertierbar ist, also $n = \text{Rang}(A) = \text{Rang}(SA)$ muß diese Zeilenstufenform wie

$$B = \begin{pmatrix} b_{11} & * & * & * \\ 0 & b_{22} & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & b_{nn} \end{pmatrix}$$

mit $b_{11} \neq 0$,..., $b_{nn} \neq 0$ aussehen. Es gibt also Elementarmatrizen $B_1, ..., B_r$ mit $B = B_r \cdots B_1 \cdot A$. Nun werden die Elemente $b_{jn}$, $j = 1, ..., n-1$ der n-ten Spalte durch Addition von $-\frac{b_{jn}}{b_{nn}}$ mal der n-ten Zeile eliminiert und wir bekommen

$$\begin{pmatrix} b_{11} & * & * & 0 \\ 0 & b_{22} & * & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & b_{nn} \end{pmatrix}$$

Verfahre ebenso mit den Elementen $b_{1k}, ..., b_{k-1,k}$ in der k-ten Spalte beginnend mit $k = n-1$ bis zu $k = 2$ und beachte, daß sich dabei die bereits transformierten Spaltenvektoren

$$\underline{\tilde{b}}_{k+1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_{k+1,k+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, ..., \underline{\tilde{b}}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_{nn} \end{pmatrix}$$ nicht mehr ändern. Wir bekommen damit weitere Elementarmatrizen $B_s, ..., B_{r+1}$, mit denen wir

$$B_s \cdots B_{r+1} B = \begin{pmatrix} b_{11} & & 0 \\ & \ddots & \\ 0 & & b_{nn} \end{pmatrix} = C$$

schreiben können. Durch Linksmultiplikation mit $S_i(b_{ii}^{-1})$, $i = 1, ..., n$ wird $C$ in die Einheitsmatrix $E_n$ überführt. Insgesamt erhalten wir also $E_n = S_1(b_{11}^{-1}) \cdots S_n(b_{nn}^{-1})C = S_1(b_{11}^{-1}) \cdots S_n(b_{nn}^{-1})B_s \cdots B_1 \cdot A$ oder

$$A = B_1^{-1} \cdots B_s^{-1} S_n(b_{nn}) \cdots S_1(b_{11}).$$

□

**Rezept 9.3.2** (zum Invertieren von Matrizen)**.** Ist $A \in M_n(K)$ gegeben, so schreiben wir zur Berechnung von $A^{-1}$ die Matrizen $A$ und $E_n$ nebeneinander und führen alle Zeilenumformungen, die wir mit $A$ machen, auch an $E_n$ aus. Zunächst bringt man $A$ auf Zeilenstufenform und unterscheidet dann folgende Fälle:

1. Ist die Zeilenstufenform von $A$ von der Gestalt

$$\begin{pmatrix} b_{11} & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & b_{nn} \end{pmatrix} \quad \text{mit } b_{11} \neq 0, ..., b_{nn} \neq 0, \tag{9.3.1}$$

so führt man weitere Umformungen wie eben im Beweis durch, bis man $A$ in die Einheitsmatrix überführt hat. Die aus $E_n$ umgeformte Matrix ist nun $A^{-1}$.

2. Ist die Zeilenstufenform von $A$ nicht von der Form (9.3.1), so ist $A$ nicht invertierbar.

**Beispiel 9.3.3.** $A = \begin{pmatrix} 0 & 1 & -4 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{pmatrix}$ Wir notieren links $A$, rechts $E_3$ und markieren noch die Zeilenumformungen:

$$\boxed{\begin{matrix} 0 & 1 & -4 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{matrix}} \quad \boxed{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}$$

$$\downarrow \ (I) \leftrightarrow (II)$$

$$\boxed{\begin{matrix} 1 & 2 & -1 \\ 0 & 1 & -4 \\ 1 & 1 & 2 \end{matrix}} \quad \boxed{\begin{matrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{matrix}}$$

$$\downarrow \ (III) - (I)$$

$$\boxed{\begin{matrix} 1 & 2 & -1 \\ 0 & 1 & -4 \\ 0 & -1 & 3 \end{matrix}} \quad \boxed{\begin{matrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 1 \end{matrix}}$$

$$\downarrow \ (III) + (II)$$

$$\boxed{\begin{matrix} 1 & 2 & -1 \\ 0 & 1 & -4 \\ 0 & 0 & -1 \end{matrix}} \quad \boxed{\begin{matrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{matrix}}$$

und wir sehen an dieser Stelle $A \in GL(3, \mathbb{R})$. Weitere Umformungen ergeben:

$$\downarrow \ (III) \cdot (-1)$$

$$\boxed{\begin{matrix} 1 & 2 & -1 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{matrix}} \quad \boxed{\begin{matrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 1 & -1 \end{matrix}}$$

$$\downarrow \ (II) + 4 \cdot (III)$$

$$\boxed{\begin{matrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}} \quad \boxed{\begin{matrix} 0 & 1 & 0 \\ -3 & 4 & -4 \\ -1 & 1 & -1 \end{matrix}}$$

$$\downarrow \ (I) + (III)$$

$$\boxed{\begin{matrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}} \quad \boxed{\begin{matrix} -1 & 2 & -1 \\ -3 & 4 & -4 \\ -1 & 1 & -1 \end{matrix}}$$

$$\downarrow \ (I) - 2 \cdot (II)$$

$$\boxed{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}} \quad \boxed{\begin{matrix} 5 & -6 & 7 \\ -3 & 4 & -4 \\ -1 & 1 & -1 \end{matrix}}$$

Wir lesen hier nun $A^{-1} = \begin{pmatrix} 5 & -6 & 7 \\ -3 & 4 & -4 \\ -1 & 1 & -1 \end{pmatrix}$ ab und können unsere Rechnung überprüfen, indem wir $AA^{-1}$ und $A^{-1}A$ ausrechnen.

**Beispiel 9.3.4.** Für $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ startet das Verfahren mit

$$\boxed{\begin{matrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{matrix}} \qquad \boxed{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}$$

$$\downarrow \ (III) - (I)$$

$$\boxed{\begin{matrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{matrix}} \qquad \boxed{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{matrix}}$$

$$\downarrow \ (III) + (II)$$

$$\boxed{\begin{matrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{matrix}} \qquad \boxed{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{matrix}}$$

und stoppt an dieser Stelle, da $\mathrm{Rang}(A) = 2$, also $A \notin GL(3, \mathbb{R})$.

# Kapitel 10

# Determinanten

## 10.1 Definition und elementare Eigenschaften

**Definition 10.1.1.** *Es sei $K$ ein Körper und $n \in \mathbb{N}$. Schreibe Matrizen*

$$A = (a_{ij})_{1 \le i,j \le n} = (\underline{a_1}|\cdots|\underline{a_n}) = \begin{pmatrix} \vec{a_1} \\ \vdots \\ \vec{a_n} \end{pmatrix} \in M_n(K)$$

*mit Spaltenvektoren $\underline{a_1}, ..., \underline{a_n}$ bzw. mit Zeilenvektoren $\vec{a_1}, ..., \vec{a_n}$. Eine Abbildung* $\det : M_n(K) \longrightarrow K$ *heißt* **Determinante***, wenn gilt:*

1. $\det$ *ist linear in jeder Zeile: Für alle $\lambda \in K$ und $i \in \{1, ..., n\}$ gilt:* [1]

$$\det \begin{pmatrix} \vdots \\ \lambda\vec{a_i} + \vec{a_i}' \\ \vdots \end{pmatrix} = \lambda \det \begin{pmatrix} \vdots \\ \vec{a_i} \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ \vec{a_i}' \\ \vdots \end{pmatrix}. \tag{10.1.1}$$

2. *Hat $A$ zwei gleiche Zeilen, so ist* $\det A = 0$.

3. $\det E_n = 1$.

**Bemerkung 10.1.2.** Das ist eine kurze Definition der Determinante – zB. braucht man nicht für jedes $n \in \mathbb{N}$ eine extra Definition. Es bleibt aber noch offen, ob allein durch diese drei Eigenschaften ein derartiges $\det(A) \in K$ eindeutig festgelegt wird, bzw. wie man dies explizit berechnet. Bevor wir uns im nächsten Abschnitt um Existenz und Eindeutigkeit der Determinante kümmern, wollen wir zunächst weitere Eigenschaften der Determinante aus der Definition ableiten, was dann nicht nur bei konkreten Berechnungen, sondern auch beim Beweis der Existenz und Eindeutigkeit der Determinante weiterhelfen wird.

---

[1] die Punkte $\vdots$ sollen andeuten, daß hier auf der linken und rechten Seite der Gleichung dieselben Einträge stehen; dann sieht man besser, wo sich etwas ändert

**Satz 10.1.3.** *1. Für alle $\lambda \in K$ ist*

$$\det \begin{pmatrix} \vdots \\ \lambda\vec{a}_i \\ \vdots \end{pmatrix} = \lambda \det \begin{pmatrix} \vdots \\ \vec{a}_i \\ \vdots \end{pmatrix}$$

$$\det(\lambda A) = \det \begin{pmatrix} \lambda\vec{a}_1 \\ \vdots \\ \lambda\vec{a}_n \end{pmatrix} = \lambda^n \det A$$

*2. Bei jeder Zeilenvertauschung bekommt* det *ein anderes Vorzeichen:*

$$\det \begin{pmatrix} \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_j \\ \vdots \end{pmatrix} = -\det \begin{pmatrix} \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_i \\ \vdots \end{pmatrix}$$

*3. Bei der Addition eines Vielfachen einer Zeile zu einer anderen Zeile ändert sich die Determinante nicht:*

$$\det \begin{pmatrix} \vdots \\ \vec{a}_i + \lambda\vec{a}_j \\ \vdots \\ \vec{a}_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_j \\ \vdots \end{pmatrix}.$$

*4. Die Determinante einer oberen Dreiecksmatrix ist das Produkt ihrer Diagonalelemente:*

$$\det \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = \lambda_1 \cdots \lambda_n$$

*5. Sind $l, m, n \in \mathbb{N}$ mit $l + m = n$, $A \in M_l(K)$, $B \in M_m(K)$ und $C \in M(l \times m, K)$, dann gilt:*

$$\det \left( \begin{array}{c|c} A & C \\ \hline 0 & B \end{array} \right) = (\det A)(\det B)$$

*6. Für die kanonischen Basisvektoren des $K^n$ – geschrieben als Zeilenvektoren $\vec{e}_1, ..., \vec{e}_n$– und jedes $\pi \in \mathcal{S}_n$ gilt:*

$$\det \begin{pmatrix} \vec{e}_{\pi(1)} \\ \vdots \\ \vec{e}_{\pi(n)} \end{pmatrix} = \operatorname{sign}(\pi)$$

*Beweis:* $\det\begin{pmatrix}\vdots\\ \lambda\vec{a}_i\\ \vdots\end{pmatrix}=\lambda\det\begin{pmatrix}\vdots\\ \vec{a}_i\\ \vdots\end{pmatrix}$ ist ein Spezialfall von Eigenschaft 1 aus Definition 10.1.1 für $\vec{a}_i'=\vec{0}$. Wendet man dieses Ergebnis n-mal an, so folgt:

$$\det(\lambda A)=\det\begin{pmatrix}\lambda\vec{a}_1\\ \lambda\vec{a}_2\\ \vdots\\ \lambda\vec{a}_{n-1}\\ \lambda\vec{a}_n\end{pmatrix}=\lambda\det\begin{pmatrix}\vec{a}_1\\ \lambda\vec{a}_2\\ \vdots\\ \lambda\vec{a}_{n-1}\\ \lambda\vec{a}_n\end{pmatrix}=...=\lambda^n\det\begin{pmatrix}\vec{a}_1\\ \vec{a}_2\\ \vdots\\ \vec{a}_{n-1}\\ \vec{a}_n\end{pmatrix}$$

Unter mehrmaligem Anwenden von Eigenschaft 2 aus Definition 10.1.1 können wir 0 als Determinante einer Matrix mit zwei gleichen Zeilen schreiben; es folgt also zusammen mit der Linearität von det in jeder Zeile:

$$\begin{aligned}
&\det\begin{pmatrix}\vdots\\ \vec{a}_i\\ \vdots\\ \vec{a}_j\\ \vdots\end{pmatrix}+\det\begin{pmatrix}\vdots\\ \vec{a}_j\\ \vdots\\ \vec{a}_i\\ \vdots\end{pmatrix}=\\
=\;&\det\begin{pmatrix}\vdots\\ \vec{a}_i\\ \vdots\\ \vec{a}_i\\ \vdots\end{pmatrix}+\det\begin{pmatrix}\vdots\\ \vec{a}_i\\ \vdots\\ \vec{a}_j\\ \vdots\end{pmatrix}+\det\begin{pmatrix}\vdots\\ \vec{a}_j\\ \vdots\\ \vec{a}_i\\ \vdots\end{pmatrix}+\det\begin{pmatrix}\vdots\\ \vec{a}_j\\ \vdots\\ \vec{a}_j\\ \vdots\end{pmatrix}=\\
=\;&\det\begin{pmatrix}\vdots\\ \vec{a}_i\\ \vdots\\ \vec{a}_i+\vec{a}_j\\ \vdots\end{pmatrix}+\det\begin{pmatrix}\vdots\\ \vec{a}_j\\ \vdots\\ \vec{a}_i+\vec{a}_j\\ \vdots\end{pmatrix}=\det\begin{pmatrix}\vdots\\ \vec{a}_i+\vec{a}_j\\ \vdots\\ \vec{a}_i+\vec{a}_j\\ \vdots\end{pmatrix}=0
\end{aligned}$$

Zum Nachweis von Aussage 3 verwende man wieder diese beiden Eigenschaften der Determinante:

$$\det\begin{pmatrix}\vdots\\ \vec{a}_i+\lambda\vec{a}_j\\ \vdots\\ \vec{a}_j\\ \vdots\end{pmatrix}=\det\begin{pmatrix}\vdots\\ \vec{a}_i\\ \vdots\\ \vec{a}_j\\ \vdots\end{pmatrix}+\lambda\det\begin{pmatrix}\vdots\\ \vec{a}_j\\ \vdots\\ \vec{a}_j\\ \vdots\end{pmatrix}=\det\begin{pmatrix}\vdots\\ \vec{a}_i\\ \vdots\\ \vec{a}_j\\ \vdots\end{pmatrix}.$$

Zur Berechnung der Determinante einer oberen Dreiecksmatrix unterscheiden wir zwei Fälle:

a) Es gibt ein Diagonalelement $\lambda_k = 0$ ($k \in \{1, ..., n\}$). Dann kann diese Matrix durch elementare Zeilenumformungen wie dem Vertauschen von Zeilen und Addition eines Vielfachen einer anderen Zeile auf die Form

$$B = \begin{pmatrix} \lambda_1 & * & \cdots & \cdots & \cdots & \cdots & \cdots & * \\ 0 & \lambda_2 & \ddots & & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & & \ddots & \lambda_{k-1} & \ddots & & & \vdots \\ \vdots & & & \ddots & \tilde{\lambda}_k & \ddots & & \vdots \\ \vdots & & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & & \ddots & \tilde{\lambda}_{n-1} & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{pmatrix}$$

gebracht werden. Die Determinante ändert dabei – nach Teil 1,2 und 3 dieses Satzes – höchstens das Vorzeichen und nach Teil 1 dieses Satzes angewandt auf den letzten Zeilenvektor $\vec{0} = 0 \cdot \vec{0}$ von $B$, ist $\det B = 0 = \lambda_1 \cdots \lambda_n$.

b) Ist $\lambda_k \neq 0$ für alle $k = 1, ..., n$, so ergibt die Anwendung von Teil 1 dieses Satzes

$$\det \begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & b_{n-1,n} \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = \lambda_1 \cdots \lambda_n \det \begin{pmatrix} 1 & \frac{b_{12}}{\lambda_1} & \cdots & \frac{b_{1n}}{\lambda_1} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{b_{n-1,n}}{\lambda_{n-1}} \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Nun überführt man wieder die Matrix $\begin{pmatrix} 1 & \frac{b_{12}}{\lambda_1} & \cdots & \frac{b_{1n}}{\lambda_1} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{b_{n-1,n}}{\lambda_{n-1}} \\ 0 & \cdots & 0 & 1 \end{pmatrix}$ durch Addition von Vielfachen von Zeilen in die Einheitsmatrix. Der Eintrag $\frac{b_{n-1,n}}{\lambda_{n-1}}$ wird beispielsweise durch Subtraktion von $\frac{b_{n-1,n}}{\lambda_{n-1}}$ mal der n-ten Zeile von der (n-1)-ten Zeile eliminiert:

$$\begin{pmatrix} 1 & \frac{b_{1,2}}{\lambda_1} & \cdots & \frac{b_{1,n-2}}{\lambda_1} & \frac{b_{1,n-1}}{\lambda_1} & \frac{b_{1,n}}{\lambda_1} \\ 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & \ddots & & \\ \vdots & & & 1 & \frac{b_{n-2,n-1}}{\lambda_{n-2}} & \frac{b_{n-2,n}}{\lambda_{n-2}} \\ \vdots & & & & 1 & \frac{b_{n-1,n}}{\lambda_{n-1}} \\ 0 & \cdots & & & & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & \frac{b_{1,2}}{\lambda_1} & \cdots & \frac{b_{1,n-2}}{\lambda_1} & \frac{b_{1,n-1}}{\lambda_1} & \frac{b_{1,n}}{\lambda_1} \\ 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & \ddots & & \\ \vdots & & & 1 & \frac{b_{n-2,n-1}}{\lambda_{n-2}} & \frac{b_{n-2,n}}{\lambda_{n-2}} \\ \vdots & & & & 1 & 0 \\ 0 & \cdots & & & & 1 \end{pmatrix}$$

Anschließend wird der Eintrag $\frac{b_{n-2,n}}{\lambda_{n-2}}$ in der $(n-2)-$ten Zeile, $n-$te Spalte durch Subtraktion von $\frac{b_{n-2,n}}{\lambda_{n-2}}$ mal der $n-$ten Zeile eliminiert. Durch wiederholte Anwendung dieses Verfahrens kommen wir zu

$$\begin{pmatrix} 1 & \frac{b_{1,2}}{\lambda_1} & \cdots & \frac{b_{1,n-2}}{\lambda_1} & \frac{b_{1,n-1}}{\lambda_1} & 0 \\ 0 & 1 & \ddots & \vdots & \vdots & \\ \vdots & & \ddots & & & \\ \vdots & & & 1 & \frac{b_{n-2,n-1}}{\lambda_{n-2}} & 0 \\ \vdots & & & & 1 & 0 \\ 0 & \cdots & & & & 1 \end{pmatrix}$$

Durch Subtraktion eines geeigneten Vielfachen der $(n-1)-$ten Zeile, lassen sich die Einträge in der $(n-1)-$ten Spalte über der Diagonale sukzessive eliminieren und wir erhalten

$$\begin{pmatrix} 1 & \frac{b_{1,2}}{\lambda_1} & \cdots & \frac{b_{1,n-2}}{\lambda_1} & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots & \\ \vdots & & \ddots & & & \\ \vdots & & & 1 & 0 & 0 \\ \vdots & & & & 1 & 0 \\ 0 & \cdots & & & & 1 \end{pmatrix}$$

Macht man das so auch mit den Spalten $n-2$ bis 2, so gelangen wir zur Einheitsmatrix $E_n$. Nach Teil 3 dieses Satzes ändert sich dabei die Determinate nicht, also gilt

$$\det \begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & b_{n-1,n} \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = \lambda_1 \cdots \lambda_n \det E_n = \lambda_1 \cdots \lambda_n$$

gemäß der Normierungsbedingung $\det E_n = 1$.

Bringe durch Addition eines Vielfachen einer Zeile zu einer anderen Zeile und Zeilenvertauschungen die Matrix $(A|C) \in M(l \times (m+l), K)$ auf eine Zeilenstufenform $(A'|C')$. Durch dieselben Zeilenumformungen - mit $k_1$ Zeilenvertauschungen - wird aus $\left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right)$ die Matrix $\left(\begin{array}{c|c} A' & C' \\ \hline 0 & B \end{array}\right)$. Beachte, daß diese elementaren Zeilenumformungen ausschießlich auf die oberen $l$ Zeilen wirken und daher der untere Block $(0|B)$ dabei nicht verändert wird. Bringe ebenso $B$ durch elementare Zeilenumformungen, die nur auf die unteren $m$ Zeilen wirken in eine Zeilenstufenform $B'$ - mit $k_2$ Zeilenvertauschungen. Da $A'$ und $B'$ als Zeilenstufenformen von quadratischen Matrizen obere Dreiecksmatrizen sind (eventuell mit Nullen auf der Diagonalen) ist auch $\left(\begin{array}{c|c} A' & C' \\ \hline 0 & B' \end{array}\right)$ eine obere Dreiecksmatrix. Nach den eben gezeigten Regeln ist $\det A = (-1)^{k_1} \det(A')$, $\det B = (-1)^{k_2} \det(B')$ und $\det \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right) =$

$(-1)^{k_1+k_2} \det \left(\begin{array}{c|c} A' & C' \\ \hline 0 & B' \end{array}\right)$. Nach der Rechenregel für die Determinante oberer Dreiecksmatizen ist $\det \left(\begin{array}{c|c} A' & C' \\ \hline 0 & B' \end{array}\right) = \det(A') \det(B')$, also nach Berücksichtigungen der Vorzeichen $\det \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right) = \det(A) \det(B)$. Zu $\pi \in \mathcal{S}_n$ finden wir nach Lemma 3.1.15 Transpositionen $\tau_1, ..., \tau_r$ mit $\pi = \tau_1 \circ ... \circ \tau_r$, also ist $\operatorname{sign}(\pi) = (-1)^r$. Daher geht $\begin{pmatrix} \vec{e}_{\pi(1)} \\ \vdots \\ \vec{e}_{\pi(n)} \end{pmatrix}$ durch $r$ Zeilenvertauschungen aus $\begin{pmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_n \end{pmatrix}$ hervor und nach Rechenregel 2 aus diesem Satz folgt:

$$\det \begin{pmatrix} \vec{e}_{\pi(1)} \\ \vdots \\ \vec{e}_{\pi(n)} \end{pmatrix} = (-1)^r \det \begin{pmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_n \end{pmatrix} = (-1)^r \det E_n = \operatorname{sign}(\pi).$$

$\square$

**Rezept 10.1.4.** Ist $A \in M_n(K)$ und soll $\det(A)$ berechnet werden, so forme $A$ durch elementare Zeilenumformungen – wie dem $k$-fachen Vertauschen von Zeilen und wiederholter Addition eines Vielfachen einer Zeile zu einer anderen – in eine Matrix $B$ von Zeilenstufenform um. Nach Satz 10.1.3 gilt also $\det(A) = (-1)^k \det(B)$ und da $B$ dann eine obere Dreiecksmatrix ist (möglicherweise mit Nullen auf der Diagonale!) läßt sich $\det(B)$ wieder nach Satz 10.1.3 berechnen.

**Beispiel 10.1.5.**

a) Die Matrix $A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix}$ wird durch elementare Zeilenumformungen

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix} \xrightarrow[(II)\leftrightarrow(III')]{(I)\leftrightarrow(III)} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} \xrightarrow{(III)-2(II)} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

mit zwei Zeilenvertauschungen in eine obere Dreiecksmatrix überführt. Daher ist

$$\det \begin{pmatrix} 0 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix} = (-1)^2 \det \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix} = -1$$

b) Die Matrix $B = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 3 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ 2 & 2 & 1 & 0 \end{pmatrix}$ wird durch elementare Zeilenumformungen

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 3 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ 2 & 2 & 1 & 0 \end{pmatrix} \overset{\substack{(II)-2(I) \\ (III)-(I) \\ (IV)-2(I)}}{\longrightarrow} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & -2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & -2 & -1 & 0 \end{pmatrix} \overset{(IV)-2(II)}{\longrightarrow} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & -2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 3 & -2 \end{pmatrix}$$

$$\overset{(IV)\leftrightarrow(III)}{\longrightarrow} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & -2 & 1 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

mit einer Zeilenvertauschung in eine obere Dreiecksmatrix überführt. Daher ist

$$\det \begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 3 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ 2 & 2 & 1 & 0 \end{pmatrix} = -\det \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & -2 & 1 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 3$$

**Satz 10.1.6.** *Für $A \in M_n(K)$ sind äquivalent:*

- $\mathrm{Rang}(A) = n$
- $A \in GL(n, K)$
- $\det(A) \neq 0$

*Beweis.* Aus Satz 8.4.4 wissen wir bereits, daß $\mathrm{Rang}(A) = n$ und $A \in GL(n, K)$ äquivalent sind. Nach Satz 9.2.10 läßt sich $A$ durch das Vertauschen von Zeilen und die wiederholte Addition eines Vielfachen einer Zeile zu einer anderen Zeile auf Zeilenstufenform $B = SA$ bringen. Nach Satz 10.1.3 gilt $\det(A) = \det(B)$ oder $\det(A) = -\det(B)$, denn:

- Die Addition eines Vielfachen einer Zeile zu einer anderen Zeile ändert die Determinante nicht.
- Das Vertauschen von zwei Zeilen ändert nur das Vorzeichen der Determinante.

An einer Zeilenstufenform $B = SA \in M_n(K)$ von $A$ sieht man, daß $\mathrm{Rang}(A) = n$ äquivalent zu $B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{22} & & \\ & & \ddots & \\ & & & b_{nn} \end{pmatrix}$ mit $b_{11} \neq 0, \ldots, b_{nn} \neq 0$ ist. Aus der Rechenregel für die Determinante einer oberen Dreiecksmatrix folgt dann die Behauptung. □

**Beispiel 10.1.7.**

- $n = 1$: Eine $1 \times 1$ Matrix $A = (a) = aE_1 \in M_1(K)$ besteht aus nur einem Eintrag mit $a \in K$, dann ist

$$\det A = \det(aE_1) = a \det(E_1) = a.$$

- $n = 2$: Für eine $2 \times 2$ Matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ gilt:

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

denn im Fall $a_{11} \neq 0$ erhalten wir die Zeilenstufenform $\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} - \dfrac{a_{21}}{a_{11}} a_{12} \end{pmatrix}$ von $A$ durch Subtraktion des $\frac{a_{21}}{a_{11}}$-fachen der 1. Zeile von der 2. Zeile. Dabei ändert sich die Determinante nicht, also ist $\det A = a_{11}(a_{22} - \frac{a_{21}}{a_{11}} a_{12}) = a_{11}a_{22} - a_{12}a_{21}$. Im Fall $a_{11} = 0$ vertauschen wir die 1. und 2. Zeile, dann ist

$$\det A = \det \begin{pmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = -\det \begin{pmatrix} a_{21} & a_{22} \\ 0 & a_{12} \end{pmatrix} = -a_{21}a_{12}.$$

**Satz 10.1.8.** *Sind $A, B \in M_n(K)$, so gilt*

$$\det(AB) = \det(A) \cdot \det(B). \tag{10.1.2}$$

*Insbesondere ist* $\det(A^{-1}) = \dfrac{1}{\det(A)}$, *falls $A \in GL(n, K)$.*

*Beweis.* Ist Rang$A < n$ oder Rang$(B) < n$, so ist auch Rang$(AB) < n$ und nach Satz 10.1.6 lautet die Gleichung (10.1.2) dann $0 = \det(A)\det(B) = \det(AB) = 0$. [2] Es genügt also im Folgenden nur noch den Fall $A, B \in GL(n, K)$ zu betrachten; dann gibt es nach Satz 9.3.1 eine Faktorisierung $AB = B_1 \cdots B_r$ mit Elementarmatrizen $B_1, ..., B_r$. Betrachte zuerst den Fall $r = 2$:

- Ist $B_1 = S_i(\lambda) = \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& \lambda &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix}$, so ist $\det(S_i(\lambda)) = \lambda$ und beim Bilden von $S_i(\lambda)B_2$ wird gerade die $i$-te Zeile von $B_2$ mit $\lambda$ multipliziert, damit ist $\det(S_i(\lambda)B_2) = \lambda \det(B_2) = \det(S_i(\lambda))\det(B_2)$.

---

[2] Mit $\begin{array}{rcl} F_A : K^n & \to & K^n \\ \underline{x} & \mapsto & A\underline{x} \end{array}$, $\begin{array}{rcl} F_B : K^n & \to & K^n \\ \underline{x} & \mapsto & B\underline{x} \end{array}$ und $\begin{array}{rcl} F_{AB} : K^n & \to & K^n \\ \underline{x} & \mapsto & AB\underline{x} \end{array}$ gelten Rang$(A) = \dim_K(F_A(K^n))$, Rang$(B) = \dim_K(F_B(K^n))$ und Rang$(AB) = \dim_K(F_{AB}(K^n))$. Wegen $F_{AB}(K^n) = F_A(F_B(K^n)) \subseteq F_A(K^n)$ folgt Rang$(AB) \leq$ Rang$(A) < n$ im Fall von Rang$(A) < n$. Im Fall Rang$(B) < n$ ist $F_B(K^n) \subseteq K$ ein Rang$(B)-$dimensionaler Unterraum und Anwenden der Dimensionsformel auf

$$\begin{array}{rcl} F_A|_{F_B(K^n)} : F_B(K^n) & \to & K^n \\ \underline{y} & \mapsto & A\underline{y} \end{array}$$

gibt $n >$ Rang$(B) = \dim_K(F_B(K^n)) = \dim_K(F_A(F_B(K^n))) + \dim_K(\mathrm{Kern}(F_A|_{F_B(K^n)})) \geq \dim_K(F_{AB}(K^n)) =$ Rang$(AB)$.

- Ist $B_1 = Q_i^j(\lambda) = \begin{array}{cl} \begin{pmatrix} 1 &&&&&&& \\ & \ddots &&&&&& \\ && 1 &&&&& \\ &&& 1 & \cdots & \lambda && \\ &&&& \ddots & \vdots && \\ &&&&& 1 && \\ &&&&&& \ddots & \\ &&&&&&& 1 \end{pmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow \text{i-te Zeile} \\ \\ \\ \\ \\ \end{array} \end{array}$, so ist $\det(Q_i^j(\lambda)) = 1$ und beim Bilden von $Q_i^j(\lambda)B_2$ wird das $\lambda$-fache der $j$-ten Zeile zur $i$-ten Zeile addiert. Nach Satz 10.1.3 ändert sich dabei die Determinante nicht, woraus $\det(Q_i^j(\lambda)B_2) = \det(B_2) = \det(Q_i^j(\lambda))\det(B_2)$ folgt.

- Ist $B_1 = P_i^j$, so wird beim Bilden von $P_i^j B_2$ in $B_2$ gerade die $i$-te mit der $j$-ten Zeile vertauscht. Da $\det(P_i^j) = -1$ ist ändert nach Satz 10.1.3 die Determinante dabei das Vorzeichen: $\det(P_i^j B_2) = -\det(B_2) = \det(P_i^j)\det(B_2)$.

Für eine Elementarmatrix $B_1$ und beliebige Produkte $B$ von Elementarmatrizen gilt also (10.1.2). Ist $A \in GL(n, K)$, so existiert $A^{-1}$ und $AA^{-1} = E_n$, also $\det(A)\det(A^{-1}) = 1$. □

**Bemerkung 10.1.9.** Ist $V$ ein $n$-dimensionaler $K$-Vektorraum mit einer Basis $\mathcal{A}$ und $F : V \to V$ eine $K$-lineare Abbildung mit darstellender Matrix $A := M_{\mathcal{A}}^{\mathcal{A}}(F)$ bezüglich der Basis $\mathcal{A}$. Dann spricht man von $\det(A)$ auch als der **Determinante des Endomorphismus** $F$ und schreibt $\det(F)$, denn für jede andere Basis $\mathcal{B}$ von $V$ ist die darstellende Matrix $B = M_{\mathcal{B}}^{\mathcal{B}}(F)$ nach Bemerkung 8.4.6 ähnlich zu $A$, dh. es gibt ein $S \in GL(n, K)$ mit $B = SAS^{-1}$. Nach Satz 10.1.8 ist $\det(B) = \det(S)\det(A)\det(S^{-1}) = \det(A)$, also $\det(F)$ unabhängig von der Wahl der Basis.

## 10.2 Existenz und Eindeutigkeit der Determinante

**Satz 10.2.1.** *Ist $K$ ein Körper, $n \in \mathbb{N}$, dann gibt es genau eine Determinante, also genau eine Abbildung*

$$\det : M_n(K) \to K,$$

*die die Eigenschaften aus Definition 10.1.1 erfüllt. Für $A = (a_{ij})_{1 \le i,j \le n} \in M_n(K)$ gilt:*

$$\det(A) = \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)} \tag{10.2.1}$$

*Beweis.* Schreibe $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_n \end{pmatrix} = \begin{pmatrix} a_{11}\vec{e}_1 + ... + a_{1n}\vec{e}_n \\ \vdots \\ a_{n1}\vec{e}_1 + ... + a_{nn}\vec{e}_n \end{pmatrix}$, dann ergibt die Linearität in jeder Zeile:

$$\begin{aligned}
\det A &= \sum_{j_1=1}^{n} a_{1j_1} \det \begin{pmatrix} \vec{e}_{j_1} \\ \vec{a}_2 \\ \vdots \\ \vec{a}_n \end{pmatrix} = \sum_{j_1=1}^{n} a_{1j_1} \sum_{j_2=1}^{n} a_{2j_2} \det \begin{pmatrix} \vec{e}_{j_1} \\ \vec{e}_{j_2} \\ \vec{a}_3 \\ \vdots \\ \vec{a}_n \end{pmatrix} \\
&= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1j_1} \cdots a_{nj_n} \det \begin{pmatrix} \vec{e}_{j_1} \\ \vdots \\ \vec{e}_{j_n} \end{pmatrix}
\end{aligned}$$

Nur im Fall, daß $\begin{pmatrix} 1 & \cdots & n \\ j_1 & \cdots & j_n \end{pmatrix} \in \mathcal{S}_n$ ist, hat $\begin{pmatrix} \vec{e}_{j_1} \\ \vdots \\ \vec{e}_{j_n} \end{pmatrix}$ paarweise verschiedene Zeilen; nach Definition 10.1.1 und Satz 10.1.3 gilt

$$\det A = \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)}. \tag{10.2.2}$$

Das zeigt, daß die drei Forderungen in Definition 10.1.1 nur eine Möglichkeit zulassen um $\det A$ zu definieren, nämlich durch (10.2.2). Für die Existenz einer Determinante bleibt noch zu zeigen, daß die durch (10.2.2) definierte Abbildung tatsächlich die Forderungen von Definition 10.1.1 erfüllt:

- Wegen

  $$\begin{aligned}
  \det \begin{pmatrix} \vdots \\ \vec{a}_i + \lambda \vec{b}_i \\ \vdots \end{pmatrix} &= \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) a_{1\pi(1)} \cdots a_{i-1,\pi(i-1)} \big(a_{i\pi(i)} + \lambda b_{i\pi(i)}\big) a_{i+1,\pi(i+1)} \cdots a_{n\pi(n)} \\
  &= \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)} \\
  &\quad + \lambda \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) a_{1\pi(1)} \cdots a_{i-1,\pi(i-1)} b_{i\pi(i)} a_{i+1,\pi(i+1)} \cdots a_{n\pi(n)} \\
  &= \det \begin{pmatrix} \vdots \\ \vec{a}_i \\ \vdots \end{pmatrix} + \lambda \det \begin{pmatrix} \vdots \\ \vec{b}_i \\ \vdots \end{pmatrix}
  \end{aligned}$$

  definiert (10.2.2) eine Abbildung $\det : M_n(K) \to K$, die in jeder Zeile linear ist.

- In $A \in M_n(K)$ seien die beiden Zeilenvektoren $\vec{a}_k$ und $\vec{a}_l$ mit $k < l$ gleich und $\tau \in \mathcal{S}_n$ sei die Transposition, die $k$ und $l$ vertauscht. Mit $\mathcal{A}_n := \{\pi \in \mathcal{S}_n : \operatorname{sign}(\pi) = 1\}$ bilden $\mathcal{A}_n$ und $\mathcal{A}_n\tau := \{\pi\tau : \pi \in \mathcal{A}_n\}$ eine Zerlegung von $\mathcal{S}_n$, denn für jedes $\sigma \in \mathcal{S}_n$

mit $\operatorname{sign}(\sigma) = -1$ ist $\operatorname{sign}(\sigma\tau) = 1$, also $\sigma = (\sigma\tau)\tau \in \mathcal{A}_n\tau$. Da $\pi \in \mathcal{A}_n\tau$ und $\pi \circ \tau \in \mathcal{A}_n$ äquivalent sind, läßt sich die Summe $\sum\limits_{\pi\in\mathcal{S}_n}$ aufspalten:

$$\det A = \sum_{\pi\in\mathcal{A}_n} \operatorname{sign}(\pi) a_{1\pi(1)}\cdots a_{n\pi(n)} - \sum_{\pi\in\mathcal{A}_n} \operatorname{sign}(\pi) a_{1\pi(\tau(1))}\cdots a_{n\pi(\tau(n))}$$

Wegen $\vec{a}_k = \vec{a}_l$ folgt nach Definition von $\tau$:

$$\begin{aligned} & a_{1\pi(\tau(1))}\cdots a_{n\pi(\tau(n))} \\ = \;& a_{1\pi(1)}\cdots a_{k-1,\pi(k-1)}a_{k\pi(l)}a_{k+1,\pi(k+1)}\cdots a_{l-1,\pi(l-1)}a_{l\pi(k)}a_{l+1,\pi(l+1)}\cdots a_{n\pi(n)} \\ = \;& a_{1\pi(1)}\cdots a_{n\pi(n)}, \end{aligned}$$

also ist in diesem Fall $\det A = 0$.

- Mit $\delta_{ij} := \begin{cases} 1 & \text{für } 1 \le i = j \le n \\ 0 & \text{für } i \ne j \end{cases}$ und $\pi \in \mathcal{S}_n$ ist $\delta_{1\pi(1)}\cdots\delta_{n\pi(n)} = \begin{cases} 1 & \text{für } \pi = \mathrm{id} \\ 0 & \text{für } \pi \ne \mathrm{id} \end{cases}$
ist $\det E_n = \det((\delta_{ij})_{1\le i,j\le n}) = \sum\limits_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi)\delta_{1\pi(1)}\cdots\delta_{n\pi(n)} = 1.$ □

**Beispiel 10.2.2. Sarrus:** Determinanten einer 3×3 Matrix berechnen sich gemäß (10.2.2) zu

$$\begin{aligned} \det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + \\ &\quad + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \end{aligned}$$

denn

$$\begin{aligned} \operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} &= 1 \\ \operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} &= -1. \end{aligned}$$

Eine Merkhilfe für diese Formel: Füge die beiden ersten Spalten rechts noch einmal an die Matrix an:

$$\begin{array}{ccccc} & & - & - & - \\ a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\ & & + & + & + \end{array}$$

Die drei Produkte auf den Diagonalen von links oben nach rechts unten werden addiert und die Produkte auf den Diagonalen von links unten nach rechts oben werden subtrahiert. Für $4 \times 4$ Matrizen (und größere) kann ein solches Schema mit Multiplikation längs der Diagonalen nicht mehr funktionieren, da man so nur 8 Summanden und nicht $4! = 24$ Summanden wie in (10.2.1) bekommt.

**Korollar 10.2.3.** *Für jede Matrix $A \in M_n(K)$ gilt:* $\det A = \det A^T$.

*Beweis.* Ist $A = (a_{ij})_{1 \le i,j \le n}$, so gilt $A^T = (a^T_{ij})_{1 \le i,j \le n} = (a_{ji})_{1 \le i,j \le n}$, also folgt aus (10.2.2):

$$\det A^T = \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) a^T_{1\pi(1)} \cdots a^T_{n\pi(n)} = \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) a_{\pi(1)1} \cdots a_{\pi(n)n}.$$

Weil $k = \pi(l)$ und $l = \pi^{-1}(k)$ für alle $k, l \in \{1, ..., n\}$ und $\pi \in \mathcal{S}_n$ äquivalent sind, enthalten die Produkte $a_{\pi(1)1} \cdots a_{\pi(n)n}$ und $a_{1\pi^{-1}(1)} \cdots a_{n\pi^{-1}(n)}$ dieselben Faktoren. Wegen $\operatorname{sign}(\pi) = \operatorname{sign}(\pi^{-1})$ gilt also

$$\det A^T = \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) a_{\pi(1)1} \cdots a_{\pi(n)n} = \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi^{-1}) a_{1\pi^{-1}(1)} \cdots a_{n\pi^{-1}(n)} = \det A.$$

□

**Satz 10.2.4** (Entwicklungssatz für Determinanten)**.** *Es sei*
$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in M_n(K)$ *und zu $i, j \in \{1, ..., n\}$ bezeichne*

$$\tilde{A}_{ij} := \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix} \tag{10.2.3}$$

*die $(n-1) \times (n-1)$ Matrix, die aus $A$ durch Streichen der $i$-ten Zeile und $j$-ten Spalte hervorgeht. Dann gilt:*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det \tilde{A}_{ij} \qquad \textit{(Entwicklung nach der $i$-ten Zeile)} \tag{10.2.4}$$

$$= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det \tilde{A}_{ij} \qquad \textit{(Entwicklung nach der $j$-ten Spalte)} \tag{10.2.5}$$

*Beweis.* Es sei $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = (\underline{a}_1 \cdots \underline{a}_n) \in M_n(K)$ mit Spaltenvektoren geschrieben. Wegen Linearität von $\det A$ in der $j$-ten Spalte (folgt aus der Linearität in den Zeilen von $A^T$ und Korollar 10.2.3) und $\underline{a}_j = \sum_{i=1}^n a_{ij}\underline{e}_i$ (mit den Einheitsvektoren $\underline{e}_1, ..., \underline{e}_n$) gilt

$$\det(A) = \sum_{i=1}^n a_{ij} \det(\underline{a}_1, ..., \underline{a}_{j-1}, \underline{e}_i, \underline{a}_{j+1}, ...\underline{a}_n).$$

In der Matrix $(\underline{a}_1, ..., \underline{a}_{j-1}, \underline{e}_i, \underline{a}_{j+1}, ...\underline{a}_n)$ rücke $\underline{e}_i$ durch $j-1$ Vertauschungen benachbarter Spaltenvektoren ganz nach links und erhalte $B_{ij} = (\underline{e}_i, \underline{a}_1, ..., \underline{a}_{j-1}, \underline{a}_{j+1}, ...\underline{a}_n)$ und rücke dann den $i$-ten Zeilenvektor von $B_{ij}$ durch $i-1$ Zeilenvertauschungen benachbarter Zeilen ganz nach oben und erhalte

$$C_{ij} := \left(\begin{array}{c|cccccc} 1 & a_{i1} & \cdots & a_{i(j-1)} & a_{i(j+1)} & \cdots & a_{in} \\ \hline & & & & & & \\ \underline{0} & & & \tilde{A}_{ij} & & & \\ & & & & & & \end{array}\right).$$

Hierbei wurden $(j-1)+(i-1)$ Spalten- und Zeilenvertauschungen vorgenommen, daher ist $\det(\underline{a}_1, ..., \underline{a}_{j-1}, \underline{e}_i, \underline{a}_{j+1}, ...\underline{a}_n) = (-1)^{i+j}\det(C_{ij}) = (-1)^{i+j}\det(\tilde{A}_{ij})$, wobei die letzte Gleichheit aus der Determinantenformel für Blockmatrizen in Satz 10.1.3 folgt. Das zeigt die Formel (10.2.5) für die Entwicklung von $\det A$ nach der $j$-ten Spalte. Der Beweis von (10.2.4) geht analog. □

**Beispiel 10.2.5.** 1. $\det\begin{pmatrix} 5 & 2 & -3 \\ 4 & -7 & 0 \\ 1 & 8 & 1 \end{pmatrix} = 5\cdot(-7)\cdot 1 + 2\cdot 0\cdot 1 + (-3)\cdot 4\cdot 8 -$
$-1\cdot(-7)\cdot(-3) - 8\cdot 0\cdot 5 - 1\cdot 4\cdot 2 = -35 - 96 - 21 - 8 = -160$

2. Bei größeren Matrizen ist –sofern nicht in einer Zeile oder Spalte viele Einträge 0 sind – die Umformung auf eine obere Dreiecksmatrix durch Addition eines Vielfachen einer anderen Zeile die schnellere Methode zur Bestimmung der Determinante:

$$\det\begin{pmatrix} 1 & 2 & 5 & 4 & 8 \\ 2 & 4 & 15 & 8 & 16 \\ 1 & 2 & 10 & 8 & 8 \\ 2 & 4 & 15 & 12 & 24 \\ 1 & 2 & 5 & 4 & 16 \end{pmatrix} = \det\begin{pmatrix} 1 & 2 & 5 & 4 & 8 \\ 2 & 4 & 15 & 8 & 16 \\ 0 & 0 & 5 & 4 & 0 \\ 0 & 0 & 0 & 4 & 8 \\ 0 & 0 & 0 & 0 & 8 \end{pmatrix} =$$

Subtraktion der 1. von der 3. und 5. Zeile, und Subtraktion der 2. von der 4. Zeile; jetzt noch zweimal 1. Zeile von 2. Zeile subtrahieren:

$$= \det\begin{pmatrix} 1 & 2 & 5 & 4 & 8 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 4 & 0 \\ 0 & 0 & 0 & 4 & 8 \\ 0 & 0 & 0 & 0 & 8 \end{pmatrix} = 0.$$

Bei Entwicklung dieser Determinante hätten wir $5! = 120$ Produkte aus jeweils 5 Faktoren mit den richtigen Vorzeichen versehen müssen und dann summieren!

3. $\det\begin{pmatrix} 0 & 4 & 27 & 5 \\ -3 & 2 & 12 & 7 \\ 2 & 4 & 0 & 8 \\ 0 & 1 & 4 & 5 \end{pmatrix} = 3\det\begin{pmatrix} 4 & 27 & 5 \\ 4 & 0 & 8 \\ 1 & 4 & 5 \end{pmatrix} + 2\det\begin{pmatrix} 4 & 27 & 5 \\ 2 & 12 & 7 \\ 1 & 4 & 5 \end{pmatrix}$

$$= 3\cdot(27\cdot 8 + 5\cdot 4\cdot 4 - 4\cdot 4\cdot 8 - 27\cdot 4\cdot 5) + 2\cdot(4\cdot 12\cdot 5 + 27\cdot 7 +$$

$$+5\cdot 2\cdot 4 - 12\cdot 5 - 4\cdot 4\cdot 7 - 2\cdot 5\cdot 27) = -1116 + 54 = -1062.$$

**Lemma 10.2.6.** *Zu $A \in M_n(K)$ sei $\tilde{A}_{ij} \in M_{n-1}(K)$ wie in (10.2.3) die Matrix, die man durch Streichen der $i$-ten Zeile und $j$-ten Spalte aus $A$ erhält. Dann heißt*

$$\mathrm{Adj}(A) := \left(\left((-1)^{i+j}\det(\tilde{A}_{ij})\right)_{1\le i,j\le n}\right)^T \tag{10.2.6}$$

*die* **Adjunkte** *von $A$. Es gilt:*

$$A\cdot \mathrm{Adj}(A) = \mathrm{Adj}(A)\cdot A = \det(A)E_n \tag{10.2.7}$$

*Beweis.* Zu $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in M_n(K)$ und $i, j \in \{1, ..., n\}$ sei

$$\widetilde{A}_{ij} := \left(\begin{array}{ccc|ccc} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ \hline a_{i+1,1} & & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{array}\right) \in M_{n-1}(K)$$

die $(n-1)\times(n-1)-$Matix, die aus $A$ durch Streichen der $i-$ten Zeile und $j-$ten Spalte entsteht. Als Determinante einer Blockdiagonalmatix bzw. durch Tauschen der 1. Zeile mit den ersten $j-1$ Zeilen und Tauschen der 1. Spalte mit den ersten $k-1$ Spalten erhalten wir

$$\det(\widetilde{A}_{jk}) = \det\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \widetilde{A}_{jk} & \\ 0 & & & \end{array}\right)$$

$$= \quad (-1)^{j-1+k-1}\det\begin{pmatrix} a_{1,1} & \cdots & a_{1,k-1} & 0 & a_{1,k+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{j-1,1} & \cdots & a_{j-1,k-1} & 0 & a_{j-1,k+1} & \cdots & a_{j-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{j+1,1} & \cdots & a_{j+1,k-1} & 0 & a_{j+1,k+1} & \cdots & a_{j+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,k-1} & 0 & a_{n,k+1} & \cdots & a_{n,n} \end{pmatrix}$$

Durch Entwicklung dieser Determinante nach der $k$−ten Spalte folgt mit den Spaltenvektoren $\underline{a}_1, ..., \underline{a}_n \in \mathbb{R}^n$ von $A$ und den Standardbasisvektoren $\underline{e}_1, ..., \underline{e}_n \in \mathbb{R}^n$ dann

$$\det(\widetilde{A}_{jk}) = (-1)^{j+k} \det(\underline{a}_1, ..., \underline{a}_{k-1}, \underline{e}_j, \underline{a}_{k+1}, ..., \underline{a}_n). \tag{10.2.8}$$

Das Matrixprodukt $\text{Adj}(A) \cdot A$ hat nach Definition der Adjunkten und nach (10.2.8) die Einträge

$$\begin{aligned}
(\text{Adj}(A) \cdot A)_{rs} &= \sum_{t=1}^{n} (\text{Adj}(A))_{rt} a_{ts} = \sum_{t=1}^{n} (-1)^{r+t} \det(\widetilde{A}_{tr}) a_{ts} \\
&= \sum_{t=1}^{n} (-1)^{r+t}(-1)^{r+t} \det(\underline{a}_1, ..., \underline{a}_{r-1}, \underline{e}_t, \underline{a}_{r+1}, ..., \underline{a}_n) a_{ts} \\
&= \det\left(\underline{a}_1, ..., \underline{a}_{r-1}, \sum_{t=1}^{n} a_{ts}\underline{e}_t, \underline{a}_{r+1}, ..., \underline{a}_n\right) \\
&= \det(\underline{a}_1, ..., \underline{a}_{r-1}, \underline{a}_s, \underline{a}_{r+1}, ..., \underline{a}_n) \\
&= \begin{cases} \det(A) & \text{für } r = s \\ 0 & \text{für } r \neq s \end{cases}
\end{aligned}$$

denn im Fall $r \neq s$ kommt der Spaltenvektor $\underline{a}_s$ doppelt vor. Damit gilt $\text{Adj}(A) \cdot A = \det(A) E_n$. Analog folgt auch $A \cdot \text{Adj}(A) = E_n$. □

**Bemerkung 10.2.7.** Für eine $2 \times 2$ Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist $\text{Adj}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, daher ist im Fall $\det A = ad - bc \neq 0$ die inverse Matrix

$$A^{-1} = \frac{1}{\det A} \text{Adj}(A) = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

**Bemerkung 10.2.8.** Ist $R$ ein kommutativer Ring mit Einselement, so kann man die $m \times n$ Matrizen

$$M(m \times n, R) := \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} : a_{11}, ..., a_{mn} \in R \right\}$$

mit Matrixeinträgen aus $R$ betrachten. Wenn $R$ kein Körper ist, so ist auch $M(m \times n, R)$ kein Vektorraum, aber wie in Definition 8.2.3 und 8.2.5 lassen sich die Addition, Skalarmultiplikation in $M(m \times n, R)$ und Multiplikation $AB$ von Matrizen $A \in M(m \times n, R)$ und $B \in M(n \times r, R)$ definieren. $M_n(R)$ wird dadurch zu einem ia. nicht kommutativen Ring. Für $A = (a_{ij})_{1 \leq i,j \leq n} \in M_n(R)$ erhält man durch

$$\det(A) = \sum_{\pi \in \mathcal{S}_n} \text{sign}(\pi) a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)} \tag{10.2.9}$$

eine Determinantenfunktion

$$\det : M_n(R) \to R,$$

die alle Bedingungen von Definition 10.1.1 erfüllt – das sieht man aus dem Beweis von Satz 10.2.1, denn dabei wurde nirgends die spezielle Eigenschaften des Körpers $K$ benutzt, die ein kommutativer Ring $R$ nicht hat – nämlich daß jedes Element aus $K\backslash\{0\}$ invertierbar ist. Ausgehend von (10.2.9) lassen sich auch alle Aussagen von Satz 10.1.3 und $\det(AB) = \det(A)\det(B)$ für $\det : M_n(R) \to R$ beweisen. Die Frage, wann $A \in M_n(R)$ eine inverse Matrix besitzt, beantwortet Lemma 10.2.6, denn die Gleichheit (10.2.7) bleibt auch für $A \in M_n(R)$ gültig, daher besitzt $A \in M_n(R)$ genau dann eine inverse Matrix, falls $\det A$ in $R$ ein Inverses besitzt. In diesem Fall ist $A^{-1} = [\det(A)]^{-1}\mathrm{Adj}(A)$.

# Kapitel 11

# Eigenwerte und Eigenvektoren

## 11.1 Eigenwerte und Eigenvektoren

**Definition 11.1.1.** *Es sei $V$ ein $K$-Vektorraum, $\lambda \in K$ und $F : V \to V$ sei $K$-linear.*

- $v \in V\backslash\{\mathbf{0}\}$ *heißt* **Eigenvektor** *von $F$ zum* **Eigenwert** $\lambda$, *wenn*

$$F(v) = \lambda v \tag{11.1.1}$$

*ist.*

- *Ist $\lambda$ Eigenwert des Endomorphismus $F : V \to V$, so heißt*

$$\mathrm{Eig}(F, \lambda) := \{v \in V : F(v) = \lambda v\} = \mathrm{Kern}(F - \lambda \mathrm{id}_V) \tag{11.1.2}$$

**Eigenraum** *von $F$ zum Eigenwert $\lambda$.*

- *Eine $K$-lineare Abbildung $F : V \to V$ heißt* **diagonalisierbar**, *wenn es eine Basis $\mathcal{B}$ von $V$ bestehend aus Eigenvektoren gibt.*

- *Zu $A \in M_n(K)$ heißt $\underline{v} \in K^n\backslash\{\mathbf{0}\}$ mit*

$$A\underline{v} = \lambda\underline{v} \tag{11.1.3}$$

*Eigenvektor von $A$ zum Eigenwert $\lambda$.*

**Bemerkung 11.1.2.**

- $v = \mathbf{0}$ kann nach Definition kein Eigenvektor sein, aber $\lambda = 0$ kann Eigenwert sein, zB. ist $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, also ist 0 Eigenwert von $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ zum Eigenvektor $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

- $\mathrm{Eig}(F, \lambda) \subseteq V$ und $\mathrm{Eig}(A, \lambda) = \mathrm{Kern}(F_A - \lambda \mathrm{id}_{K^n}) \subseteq K^n$ sind Untervektorräume, insbesondere ist der Nullvektor $\mathbf{0}$ – obwohl kein Eigenvektor – ein Element des Eigenraums.

**Beispiel 11.1.3.**

1. $$\begin{pmatrix} -1 & -1 & 1 \\ -4 & 2 & 4 \\ -1 & 1 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$
also ist $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ ein Eigenvektor von $A = \begin{pmatrix} -1 & -1 & 1 \\ -4 & 2 & 4 \\ -1 & 1 & 5 \end{pmatrix}$ zum Eigenwert 2.

2. $$\begin{pmatrix} 1 & 3 & 1 \\ 0 & 7 & 0 \\ 2 & 5 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$
daher ist $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ Eigenvektor von $\begin{pmatrix} 1 & 3 & 1 \\ 0 & 7 & 0 \\ 2 & 5 & 2 \end{pmatrix}$ zum Eigenwert 0.

**Definition 11.1.4.** *Zu* $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in M_n(K)$ *heißt*

$$p_A(X) := \det(A - XE_n) = \det \begin{pmatrix} a_{11} - X & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - X & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} - X \end{pmatrix} \in K[X] \quad (11.1.4)$$

*das* **charakteristische Polynom** *von* $A$.

**Bemerkung 11.1.5.** Zu $A \in M_n(K)$ ist das charakteristische Polynom

$$p_A(X) = \sum_{k=0}^{n} b_k X^k$$

ein Polynom vom Grad $n$. Einige Koeffizienten von $p_A$ lassen sich sogar einfach angeben:

$$\begin{aligned} b_n &= (-1)^n \\ b_{n-1} &= (-1)^{n-1}(a_{11} + ... + a_{nn}) \\ b_0 &= \det A. \end{aligned}$$

*Beweis.* Es sei $E_n = (\delta_{ij})_{1 \leq i,j \leq n}$; nach (10.2.1) ist

$$p_A(X) = \det(A - XE_n) = \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi)(a_{1\pi(1)} - \delta_{1\pi(1)}X) \cdots (a_{n\pi(n)} - \delta_{n\pi(n)}X).$$

Ist nun $\pi \neq \mathrm{id}$, so ist $\delta_{k\pi(k)} = 0$ für mindestens zwei Indizes $k \in \{1, ..., n\}$, also ergibt $(a_{1\pi(1)} - \delta_{1\pi(1)}X) \cdots (a_{n\pi(n)} - \delta_{n\pi(n)}X)$ ein Polynom vom Grad $\leq n-2$. Daher bestimmen sich die Koeffizienten $b_n$ und $b_{n-1}$ als die beiden Koeffizienten der Monome $X^n$ bzw. $X^{n-1}$ des Polynoms

$$(a_{11} - X) \cdots (a_{nn} - X) = (-1)^n X^n + (-1)^{n-1}(a_{11} + ... + a_{nn})X^{n-1} + q(X)$$

- mit einem Polynom $q$ vom Grad $\leq n - 2$. Setze $X = 0$ ein, so folgt $b_0 = \det(A)$. □

**Lemma 11.1.6.** *Ähnliche Matrizen haben das gleiche charakteristische Polynom.*

*Beweis.* Es sei $B = SAS^{-1}$ mit $S \in GL(n, K)$. In $M_n(K[X])$ ist $SXE_nS^{-1} = XE_n$, also

$$B - XE_n = SAS^{-1} - SXE_nS^{-1} = S(A - XE_n)S^{-1}$$

und durch Anwenden der Determinante folgt aus dem Multiplikationssatz für Determinanten: $\det(B - XE_n) = \det(S)\det(A - XE_n)\det(S^{-1}) = \det(A - XE_n)$. □

**Definition 11.1.7.** *Es sei $V$ ein $K$-Vektorraum mit $\dim_K V = n < \infty$ und $F : V \longrightarrow V$ sei $K$-linear. Dann ist*

$$p_F(X) := \det(M^{\mathcal{B}}_{\mathcal{B}}(F) - XE_n) \tag{11.1.5}$$

*eine Definition für das charakteristische Polynom $p_F$ von $F$, die unabhängig von der gewählten Basis $\mathcal{B}$ von $V$ ist.*

**Satz 11.1.8.** *Es sei $V$ ein $n$-dimensionaler $K$-Vektorraum und $F : V \to V$ sei $K$-linear. Dann sind äquivalent:*

1. *$\lambda \in K$ ist ein Eigenwert von $F$.*
2. *$p_F(\lambda) = \det(M^{\mathcal{B}}_{\mathcal{B}}(F) - \lambda E_n) = 0$.*

*Beweis.* $\lambda$ ist genau dann ein Eigenwert von $F$, wenn es einen Eigenvektor $v \in V\setminus\{\mathbf{0}\}$ mit $F(v) = \lambda v$ gibt. Der Koordinatenvektor $\underline{v} \in K^n\setminus\{\mathbf{0}\}$ von $v$ bezüglich einer Basis $\mathcal{B} = (w_1, ..., w_n)$ ist eine Lösung des Gleichungssystems $A\underline{v} = \lambda\underline{v}$ für die darstellende Matrix $A := M^{\mathcal{B}}_{\mathcal{B}}(F)$. Das Gleichungssystem

$$A\underline{x} - \lambda\underline{x} = (A - \lambda E_n)\underline{x} = \underline{0}$$

besitzt immer die Lösung $\underline{x} = \underline{0}$, daher ist $\lambda$ genau dann Eigenwert von $A$, wenn $(A - \lambda E_n)\underline{x} = \underline{0}$ nicht eindeutig lösbar ist. Nach Satz 8.4.4 ist dies äquivalent dazu, daß $A - \lambda E_n$ nicht invertierbar ist, was nach Satz 10.1.6 genau dann gilt, wenn $\det(A - \lambda E_n) = 0$ ist. □

**Rezept 11.1.9** (Bestimmung von Eigenwerten und Eigenvektoren einer linearen Abbildung)**.**
Zu einer $K$-linearen Abbildung $F : V \longrightarrow V$ eines $n$-dimensionalen $K$-Vektorraums $V$ betrachte die darstellende Matrix $A = M^{\mathcal{B}}_{\mathcal{B}}(F)$ von $F$ in einer Basis $\mathcal{B} = (w_1, ..., w_n)$ von $V$. Zur Bestimmung der Eigenwerte von $A$ – die mit den Eigenwerten von $F$ übereinstimmen – muß man die Nullstellen des charakteristischen Polynoms $p_A$ von $A$ bestimmen. Ist dann $\lambda \in K$ eine der ermittelten Nullstellen, so ergeben sich die Eigenvektoren $\underline{v} \in K^n$ von $A$ zum Eigenwert $\lambda$ als die von Null verschiedenen Lösungen des homogenen linearen Gleichungssystems

$$(A - \lambda E_n)\underline{v} = \underline{0}.$$

$\underline{v}$ ist der Koordinatenvektor bezüglich $\mathcal{B}$ eines Eigenvektors $v \in V\setminus\{\mathbf{0}\}$ von $F$. Es gilt:

$$\mathrm{Eig}(F, \lambda) = \{v = v_1w_1 + ... + v_nw_n \in V : \underline{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathrm{Eig}(A, \lambda)\}$$

**Beispiel 11.1.10.** Zu $A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \in M_3(\mathbb{R})$ ist das charakteristische Polynom

$$\begin{aligned} p_A(X) &= \det \begin{pmatrix} -X & -1 & 1 \\ -3 & -2-X & 3 \\ -2 & -2 & 3-X \end{pmatrix} = \\ &= X(2+X)(3-X)+6+6-2(2+X)-6X-3(3-X) = \\ &= -X^3+X^2+X-1 = -(X-1)^2(X+1), \end{aligned}$$

also sind $\lambda_1 = 1$ und $\lambda_2 = -1$ die beiden einzig möglichen Eigenwerte von $A$. Bestimme die zugehörigen Eigenvektoren:
Für $\lambda_1 = 1$ sind die Vektoren des Eigenraums $\mathrm{Eig}(A, 1)$ Lösung von

$$(A - E_3)\underline{x} = \begin{pmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Durch elementare Zeilenumformungen (zweimal 1. Zeile von der 3. Zeile und dreimal 1. Zeile von der 2. Zeile subtrahieren) geht $A - E_3$ in die Zeilenstufenform $\begin{pmatrix} -1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ über. Deshalb ist

$$\mathrm{Eig}(A,1) = \left\{ \left( \begin{matrix} -\alpha_1 + \alpha_2 \\ \alpha_1 \\ \alpha_2 \end{matrix} \right)) \in \mathbb{R}^3 : \alpha_1, \alpha_2 \in \mathbb{R} \right\} = \mathrm{lin} \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Für $\lambda_1 = -1$ sind die Vektoren des Eigenraums $\mathrm{Eig}(A, -1)$ Lösung von

$$(A + E_3)\underline{x} = \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Umformung der Koeffizientenmatrix ergibt nun

$$\begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix} \overset{(II)+3(I),(III)+2(I)}{\longrightarrow} \begin{pmatrix} 1 & -1 & 1 \\ 0 & -4 & 6 \\ 0 & -4 & 6 \end{pmatrix} \overset{(III)-(II)}{\longrightarrow} \begin{pmatrix} 1 & -1 & 1 \\ 0 & -4 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

und daraus ergibt sich

$$\mathrm{Eig}(A,-1) = \left\{ \begin{pmatrix} \frac{1}{2}\alpha \\ \frac{3}{2}\alpha \\ \alpha \end{pmatrix} \in \mathbb{R}^3 : \alpha \in \mathbb{R} \right\} = \mathrm{lin} \left\{ \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}.$$

Wegen $\det \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} = -2 \neq 0$ haben wir mit $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}$ eine Basis von $\mathbb{R}^3$ gefunden, in der die lineare Abbildung $F_A$, die in der Standardbasis durch $A$ beschrieben wird, die Diagonalform $M_{\mathcal{B}}^{\mathcal{B}}(F_A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ annimmt.

Im letzten Beispiel konnten wir durch die Berechnung von Eigenvektoren die gegebene Matrix „diagonalisieren“. Dazu noch zwei elementare Beobachtungen:

**Lemma 11.1.11.** *Es sei* $\dim_K V = n < \infty$ *und* $F : V \longrightarrow V$ *sei* $K$*-linear, dann sind äquivalent:*

- *Es gibt eine Basis* $\mathcal{B} = (v_1, ..., v_n)$ *von* $V$ *aus Eigenvektoren von* $F$*.*
- *Es gibt eine Basis* $\mathcal{B}$ *von* $V$*, so daß die darstellende Matrix* $M_{\mathcal{B}}^{\mathcal{B}}(F)$ *eine Diagonalmatrix ist. (Die Einträge auf der Diagonalen von* $M_{\mathcal{B}}^{\mathcal{B}}(F)$ *sind gerade die Eigenwerte von* $F$*.)*

*Beweis.* In der $j$-ten Spalte jeder darstellenden Matrix $M_{\mathcal{B}}^{\mathcal{B}}(F)$ steht der Koordinatenvektor von $F(v_J)$ bezüglich der Basis $\mathcal{B}$, was zusammen mit der Einführung der Koordinatenvektoren und der Definition eines Eigenvektors die behauptete Äquivalenz zeigt. □

**Lemma 11.1.12.** *Sind* $v_1, ..., v_r$ *Eigenvektoren von* $F$ *zu paarweise verschiedenen Eigenwerten* $\lambda_1, ..., \lambda_r \in K$*, also* $F(v_k) = \lambda_k v_k$ *und* $\lambda_k \neq \lambda_l$ *für* $1 \leq k < l \leq r$*, so sind* $v_1, ..., v_r$ *linear unabhängig.*

*Beweis.*

- Induktionsanfang $r = 1$: Wegen $v_1 \neq \mathbf{0}$ ist $\{v_1\}$ ein linear unabhängiges System.
- Induktionsschritt $r - 1 \to r$: Sei nun $r \geq 2$ und die Behauptung für $r - 1$ bereits gezeigt und $v_1, ..., v_r$ Eigenvektoren zu den paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_r$. Sind nun $\alpha_1, ..., \alpha_r \in K$ gegeben mit

  $$\alpha_1 v_1 + ... + \alpha_r v_r = \mathbf{0}, \tag{11.1.6}$$

  so erhalten wir durch Multiplikation dieser Gleichung mit dem Eigenwert $\lambda_r$:

  $$\mathbf{0} = \lambda_r \cdot \mathbf{0} = \lambda_r \alpha_1 v_1 + ... + \lambda_r \alpha_r v_r \tag{11.1.7}$$

  und aus den Eigenwertgleichungen und der Linearität von $F$ folgt:

  $$\mathbf{0} = F(\mathbf{0}) = F(\alpha_1 v_1 + ... + \alpha_r v_r) = \lambda_1 \alpha_1 v_1 + ... + \lambda_r \alpha_r v_r \tag{11.1.8}$$

  Subtrahiere (11.1.8) von (11.1.7), so ergibt sich

  $$\mathbf{0} = \alpha_1(\lambda_r - \lambda_1)v_1 + ... + \alpha_{r-1}(\lambda_r - \lambda_{r-1})v_{r-1} \tag{11.1.9}$$

  Nach Induktionsannahme sind $v_1, ..., v_{r-1}$ linear unabhängig. Weil die Eigenwerte $\lambda_1, ..., \lambda_r$ paarweise verschieden sind, ist $\lambda_r - \lambda_1 \neq 0$,...,$\lambda_r - \lambda_{r-1} \neq 0$, so folgt $\alpha_1 = ... = \alpha_{r-1} = 0$ aus der linearen Unabhängigkeit von $v_1, ..., v_{r-1}$ und (11.1.9). Eingesetzt in (11.1.6) erhalten wir $\alpha_r = 0$ (denn $v_r \neq 0$), was die lineare Unabhängigkeit von $v_1, ..., v_r$ zeigt und den Induktionsbeweis beendet.

□

**Bemerkung 11.1.13.** Bei der Bestimmung der Nullstellen des charakteristischen Polynoms $p_A$ einer Matrix $A \in M_n(K)$ ist zu beachten:

1. Nach dem Fundamentalsatz der Algebra besitzt jedes Polynom mit komplexen Koeffizienten vom Grad $n$ genau $n$ komplexe Nullstellen, also gibt es für jedes charakteristische Polynom $p_A$ einer Matrix $A \in M_n(\mathbb{C})$ eine Zerlegung der Form

$$p_A(X) = (\lambda_1 - X)^{\mu_1} \cdots (\lambda_r - X)^{\mu_r}$$

mit paarweise verschiedenen $\lambda_1, ..., \lambda_r \in \mathbb{C}$ , $r \in \mathbb{N}$, $r \leq n$, $\mu_1, ..., \mu_r \in \mathbb{N}$ und $\mu_1 + ... + \mu_r = n$.

2. Ist $A \in M_n(\mathbb{R})$, so besitzt $p_A$ – aufgefaßt als $p_A \in \mathbb{C}[X]$ –zwar immer noch $n$ komplexe Nullstellen, aber es müssen nicht alle diese Nullstellen reelle Zahlen sein, zB. bei $A = \begin{pmatrix} -1 & -1 \\ 2 & 1 \end{pmatrix}$ ist

$$p_A(X) = \det(A - XE_2) = \det \begin{pmatrix} -1 - X & -1 \\ 2 & 1 - X \end{pmatrix} = (X+1)(X-1)+2 = X^2+1$$

und $p_A$ besitzt zwar die beiden komplexen Nullstellen $i$ und $-i$, aber keine reelle Nullstelle, also hat die Matrix $A \in M_2(\mathbb{R})$ keinen reellen Eigenwert und keinen Eigenvektor in $\mathbb{R}^2$. Falls wir $A$ als darstellende Matrix der $\mathbb{C}$-linearen Abbildung $\begin{array}{rccc} F_A : & \mathbb{C}^2 & \longrightarrow & \mathbb{C}^2 \\ & \underline{x} & \mapsto & A\underline{x} \end{array}$ lesen, dann hat $A$ die beiden Eigenwerte $i, -i \in \mathbb{C}$ und Eigenvektoren in $\mathbb{C}^2$; kann man ganz "rezeptmäßig" bestimmen, nämlich zu:

$$\begin{aligned} \mathrm{Eig}_{\mathbb{C}}(A, i) &= \{(\frac{-\alpha}{1+i}, \alpha) : \alpha \in \mathbb{C}\} \\ \mathrm{Eig}_{\mathbb{C}}(A, -i) &= \{(\frac{\alpha}{i-1}, \alpha) : \alpha \in \mathbb{C}\} \end{aligned}$$

## 11.2 Schursche Normalform

**Satz 11.2.1.** *Es sei $V$ ein $n$-dimensionaler $K$-Vektorraum und $F : V \to V$ sei $K$-linear. Dann sind äquivalent:*

*a) Es gibt eine Basis $\mathcal{B}$ von $V$, so daß*

$$M^{\mathcal{B}}_{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & * & * \\ & \ddots & * \\ 0 & 0 & \lambda_n \end{pmatrix}$$

*eine rechte obere Dreiecksmatrix ist.*

*b) Das charakteristische Polynom $p_F$ von $F$ hat die Form*

$$p_F(X) = (\lambda_1 - X) \cdots (\lambda_n - X)$$

*mit $\lambda_1, ..., \lambda_n \in K$ (nicht notwendig paarweise verschieden).*

*In diesem Fall nennt man $M^{\mathcal{B}}_{\mathcal{B}}(F)$ eine Darstellung von $F$ in* **Schurscher Normalform**. *Die Einträge $\lambda_1, ..., \lambda_n$ auf der Diagonalen von $M^{\mathcal{B}}_{\mathcal{B}}(F)$ sind gerade die Eigenwerte von $F$.*

*Beweis.* Die Folgerung $a) \Rightarrow b)$ folgt aus der Definition des charakteristischen Polynoms und der Rechenregel für die Determinante einer oberen Dreiecksmatrix. Beweis von $b) \Rightarrow a)$ durch Induktion nach $n$:

- Induktionsanfang $n = 1$: Aus $p_F(X) = \lambda_1 - X$ folgt nach Satz 11.1.8 die Existenz des Eigenwertes $\lambda_1 \in K$ und eines Eigenvektors $v \neq \mathbf{0}$. Wegen $1 = n = \dim_K(V)$ bildet $\mathcal{B} = \{v\}$ eine Basis von $V$ und $M_{\mathcal{B}}^{\mathcal{B}}(F) = (\lambda_1) \in M_1(K)$ ist rechte obere Dreiecksmatrix.

- Induktionsschritt $n - 1 \to n$: Es sei $n = \dim_K(V)$, $F : V \to V$ sei $K$-linear und $p_F(X) = (\lambda_1 - X) \cdots (\lambda_n - X)$. Dann ist $\lambda_1$ Eigenwert von $F$ und wir können einen Eigenvektor $v_1 \neq \mathbf{0}$ wählen: $F(v_1) = \lambda_1 v_1$. Ergänze $\{v_1\}$ zu einer Basis $\mathcal{C} = \{v_1, ..., v_n\}$ von $V$ und betrachte $W := \text{lin}\{v_2, ..., v_n\}$, dann ist $V = Kv_1 \oplus W$, daher hat für $w \in W$ das Bild $F(w)$ eine eindeutige Zerlegung der Form $F(w) = \alpha(w)v_1 + G(w)$ mit $\alpha(w) \in K$ und $G(w) \in W$. Offenbar sind $\begin{array}{rcl} \alpha : W & \to & K \\ w & \mapsto & \alpha(w) \end{array}$ und $\begin{array}{rcl} G : W & \to & W \\ w & \mapsto & G(w) \end{array}$ beide $K$-linear. Nach der Rechenregel für die Determinante einer Blockdiagonalmatrix ist $p_G(X) = (\lambda_2 - X) \cdots (\lambda_n - X)$, also gibt es nach Induktionsvoraussetzung eine Basis $\mathcal{B}_W = \{w_2, ..., w_n\}$ von $W$, so daß $M_{\mathcal{B}_W}^{\mathcal{B}_W}(G)$ eine rechte obere Dreiecksmatrix ist. Die darstellende Matrix von $F$ bezüglich der Basis $\mathcal{B} = \{v_1, w_2, ..., w_n\}$ hat daher die Form

$$M_{\mathcal{B}}^{\mathcal{B}}(F) = \left( \begin{array}{c|ccc} \lambda_1 & \alpha(w_2) & \cdots & \alpha(w_n) \\ \hline & & & \\ \underline{0} & & M_{\mathcal{B}_W}^{\mathcal{B}_W}(G) & \\ & & & \end{array} \right)$$

einer oberen Dreiecksmatrix

□

## 11.3 Diagonalisierbare Endomorphismen und die Jordan Normalform

Wir betrachten in diesem Abschnitt die Körper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, dort ist laut Fundamentalsatz der Algebra die Frage nach den Nullstellen von Polynomen weitestgehend geklärt. Wir untersuchen die Frage, ob und wann man die darstellende Matrix $A = M_{\mathcal{B}}^{\mathcal{B}}(F)$ einer gegebenen linearen Abbildung $F : V \to V$ durch geschickte Wahl der Basis $\mathcal{B}$ in Diagonalform oder wenigstens in Jordan Normalform bringen kann:

**Definition 11.3.1.** *Eine Matrix $J \in M_d(\mathbb{K})$ ist in* **Jordan Normalform***, wenn es $k \leq d$, $\alpha_1, ..., \alpha_k \in \mathbb{K}$ und $d_1, ..., d_k \in \mathbb{N}$, mit $d_1 + ... + d_k = d$ und* **Jordankastl** *$Q_1, ..., Q_k$ mit*

*$Q_j = (\alpha_j)$ falls $d_j = 1$ und*

$$Q_j = \begin{pmatrix} \alpha_j & 1 & & & 0 \\ & \alpha_j & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & \alpha_j \end{pmatrix} \in M_{d_j}(\mathbb{K}),\ \text{falls } d_j > 1 \tag{11.3.1}$$

*gibt, so daß*

$$J = \begin{pmatrix} \boxed{Q_1} & & & \\ & \boxed{Q_2} & & \\ & & \ddots & \\ & & & \boxed{Q_k} \end{pmatrix} \tag{11.3.2}$$

*in Blockdiagonalform ist.*

**Definition 11.3.2.** *Es sei $V$ ein $d$-dimensionaler $\mathbb{K}$-Vektorraum und $F : V \to V$ sei $\mathbb{K}$-linear mit charakteristischem Polynom*

$$p_F(X) = (\alpha_1 - X)^{a_1} \cdots (\alpha_m - X)^{a_m} \tag{11.3.3}$$

*mit paarweise verschiedenen Eigenwerten $\alpha_1, ..., \alpha_m \in \mathbb{K}$ und $a_1, ..., a_m \in \mathbb{N}$ mit $a_1 + ... + a_m = d$. Dann heißt*

- *$a_j$ für $j = 1, ..., m$ die* **algebraische Vielfachheit** *des Eigenwerts $\alpha_j$.*
- *$g_j := \dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j))$ die* **geometrische Vielfachheit** *des Eigenwerts $\alpha_j$.*
- *Zu $l \in \mathbb{N}$ sei*

  $$(F - \alpha_j \mathrm{id}_V)^l := \underbrace{(F - \alpha_j \mathrm{id}_V) \circ \cdots \circ (F - \alpha_j \mathrm{id}_V)}_{l-mal}$$

  *die lineare Abbildung $(F - \alpha_j \mathrm{id}_V)^l : V \to V$, die durch die Komposition von $l$ Kopien der linearen Abbildung $F - \alpha_j \mathrm{id}_V : V \to V$ entsteht, dann heißt*

  $$\mathrm{Eig}(F, \alpha_j, l) := \{v \in V : (F - \alpha_j \mathrm{id}_V)^l v = \mathbf{0}\} = \mathrm{Kern}\left[(F - \alpha_j \mathrm{id}_V)^l\right] \tag{11.3.4}$$

  *der* **verallgemeinerte Eigenraum $l$-ter Stufe zum Eigenwert** *$\alpha_j$.*

**Lemma 11.3.3.** *Es sei $V$ ein endlichdimensionaler $\mathbb{K}$-Vektorraum, $F : V \to V$ eine $\mathbb{K}$-lineare Abbildung und $\alpha_j \in \mathbb{K}$ ein Eigenwert von $F$, dann gilt für die geometrische Vielfachheit $g_j$ und die algebraische Vielfachheit $a_j$ von $\alpha_j$:*

$$1 \leq g_j \leq a_j. \tag{11.3.5}$$

*Beweis.* Da $\alpha_j$ Eigenwert von $F$ ist, existiert ein Eigenvektor $v_j \neq \mathbf{0}$ von $F$ zum Eigenwert $\alpha_j$, daher ist $\{\mathbf{0}\} \neq \mathrm{Eig}(F, \alpha_j)$, also $g_j := \dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j)) \geq 1$. Es sei nun $\{v_1, ..., v_{g_j}\}$ eine Basis von $\mathrm{Eig}(F, \alpha_j)$. Wir ergänzen diese zu einer Basis $\mathcal{B} = \{v_1, ..., v_{g_j}, ..., v_d\}$ von $V$. Dann ist

$$M_{\mathcal{B}}^{\mathcal{B}}(F) = \left(\begin{array}{ccc|c} \alpha_j & & & \\ & \ddots & & A \\ & & \alpha_j & \\ \hline & 0 & & B \end{array}\right)$$

mit $B \in M_{d-g_j}(\mathbb{K})$ und $A \in M(g_j \times (d - g_j), \mathbb{K})$. Nach der Determinantenregel für Blockmatrizen ist $p_F = \det(M_{\mathcal{B}}^{\mathcal{B}}(F) - XE_d) = \det(B - XE_{d-g_j})(\alpha_j - X)^{g_j}$ und damit $g_j \leq a_j$. $\square$

**Lemma 11.3.4.** *Es sei $V$ ein $d$-dimensionaler $\mathbb{K}$-Vektorraum, $F : V \to V$ sei $\mathbb{K}$-linear und $\alpha_j$ ein Eigenwert von $F$. Dann gibt es ein $\nu_j \in \mathbb{N}$ mit*

$$\{\mathbf{0}\} \subsetneq \mathrm{Eig}(F, \alpha_j) \subsetneq \mathrm{Eig}(F, \alpha_j, 2) \subsetneq ... \subsetneq \mathrm{Eig}(F, \alpha_j, \nu_j) = \mathrm{Eig}(F, \alpha_j, l) \tag{11.3.6}$$

*für alle $l \geq \nu_j$.*

*Beweis.* Da $\alpha_j$ ein Eigenwert von $F$ ist, folgt $\{\mathbf{0}\} \subsetneq \mathrm{Eig}(F, \alpha_j)$. Ist $l \in \mathbb{N}$, $u \in \mathrm{Eig}(F, \alpha_j, l)$, dh. $(F - \alpha_j \mathrm{id}_V)^l u = \mathbf{0}$, so folgt $(F - \alpha_j \mathrm{id}_V)^{l+1} u = \mathbf{0}$ durch erneute Anwendung von $F - \alpha_j \mathrm{id}_V$, dh.

$$\mathrm{Eig}(F, \alpha_j, l) \subseteq \mathrm{Eig}(F, \alpha_j, l + 1).$$

Gilt $\mathrm{Eig}(F, \alpha_j, l) \subsetneq \mathrm{Eig}(F, \alpha_j, l+1)$, folgt $\dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j, l+1)) \geq 1 + \dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j, l))$, was wegen $d = \dim_{\mathbb{K}}(V) \geq \dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j, l))$ nur endlich viele Male passieren kann. Deshalb gibt es ein $\nu_j \in \mathbb{N}$ mit

$$\{\mathbf{0}\} \neq \mathrm{Eig}(F, \alpha_j) \subseteq \mathrm{Eig}(F, \alpha_j, 2) \subseteq ... \subseteq \mathrm{Eig}(F, \alpha_j, \nu_j) = \mathrm{Eig}(F, \alpha_j, l)$$

für alle $l \geq \nu_j$. Es sei nun $\mathrm{Eig}(F, \alpha_j, l) = \mathrm{Eig}(F, \alpha_j, l + 1)$ und $u \in \mathrm{Eig}(F, \alpha_j, l + 2)$, dh. $(F - \alpha_j \mathrm{id}_V)^{l+2} u = \mathbf{0}$ oder $v := (F - \alpha_j \mathrm{id}_V)u \in \mathrm{Eig}(F, \alpha_j, l + 1) = \mathrm{Eig}(F, \alpha_j, l)$. Dann ist $\mathbf{0} = (F - \alpha_j \mathrm{id}_V)^l v = (F - \alpha_j \mathrm{id}_V)^{l+1} u$, dh. $\mathrm{Eig}(F, \alpha_j, l + 2) \subseteq \mathrm{Eig}(F, \alpha_j, l + 1)$ und das zeigt (11.3.6). $\square$

**Definition 11.3.5.** *In der Situation von Lemma 11.3.4 heißt*

$$\mathcal{H}(\alpha_j) := \mathrm{Eig}(F, \alpha_j, \nu_j) \tag{11.3.7}$$

*der* **Hauptraum** *von $F$ zum Eigenwert $\alpha_j$.*

**Lemma 11.3.6.** *Es sei $V$ ein $d$-dimensionaler $\mathbb{K}$-Vektorraum und $F : V \to V$ sei $\mathbb{K}$-linear und*

$$p_F(X) = (\alpha_1 - X)^{a_1} \cdots (\alpha_m - X)^{a_m} \tag{11.3.8}$$

*sei das charakteristische Polynom von $F$ mit paarweise verschiedenen $\alpha_1, ..., \alpha_m \in \mathbb{K}$ und $a_1, ..., a_m \in \mathbb{N}$ mit $a_1 + ... + a_m = d$. Dann gilt:*

- $F(\mathcal{H}(\alpha_j)) \subseteq \mathcal{H}(\alpha_j)$
- $\dim(\mathcal{H}(\alpha_j)) = a_j$ *für alle* $j = 1, ..., m.$
- *Für jedes* $j \in \{1, ..., m\}$ *ist*

$$V = \text{Kern}\left[(F - \alpha_j \text{id}_V)^{\nu_j}\right] \oplus \left[(F - \alpha_j \text{id}_V)^{\nu_j}\right](V) \tag{11.3.9}$$

- $V = \mathcal{H}(\alpha_1) \oplus ... \oplus \mathcal{H}(\alpha_m).$

*Beweis.*

- Wegen $F \circ \text{id}_V = \text{id}_V \circ F$ folgt für alle $l \in \mathbb{N}$ aus dem binomischen Lehrsatz - im Ring $\text{End}_{\mathbb{K}}(V) = \{F : V \to V : F \text{ ist } \mathbb{K}\text{-linear}\}$ angewandt auf die kommutierenden Elemente $F$ und $\alpha_j \text{id}_V$:

$$(F - \alpha_j \text{id}_V)^l \circ F = \sum_{k=0}^{l} \binom{l}{k} F^{k+1}(-\alpha_j \text{id}_V)^{l-k} = F \circ (F - \alpha_j \text{id}_V)^l.$$

  Zu $v \in \mathcal{H}(\alpha_j) = \text{Eig}(F, \alpha_j, \nu_j)$ ist also

$$(F - \alpha_j \text{id}_V)^{\nu_j}(Fv) = F((F - \alpha_j \text{id}_V)^{\nu_j} v) = F(\mathbf{0}) = \mathbf{0}$$

  dh. $F(\mathcal{H}(\alpha_j)) \subseteq \mathcal{H}(\alpha_j)$.

- Nach Satz 11.2.1 können wir aufgrund von (11.3.8) eine Basis $\mathcal{B}_j$ von $V$ wählen, so daß die darstellende Matrix

$$A := M^{\mathcal{B}_j}_{\mathcal{B}_j}(F) = \left(\begin{array}{c|c} B & M \\ \hline 0 & C \end{array}\right)$$

  von Schurscher Normalform ist, also $B = \begin{pmatrix} \alpha_j & * & * \\ & \ddots & * \\ & & \alpha_j \end{pmatrix} \in M_{a_j}(\mathbb{K})$ und

$$C = \begin{pmatrix} \alpha_1 & * & & & & & & & & & \\ & \ddots & * & & & & & & & & \\ & & \alpha_1 & * & & & & & & & \\ & & & \ddots & * & & & & & & \\ & & & & \alpha_{j-1} & * & & & & & \\ & & & & & \ddots & * & & & & \\ & & & & & & \alpha_{j-1} & * & & & \\ & & & & & & & \alpha_{j+1} & * & & \\ & & & & & & & & \ddots & & \end{pmatrix} \in M_{d-a_j}(\mathbb{K}) \text{ rechte obere}$$

  Dreiecksmatrizen sind und $M \in M(a_j \times (d - a_j), \mathbb{K})$. Dann ist

$$A - \alpha_j E_d = \left(\begin{array}{c|c} B - \alpha_j E_{a_j} & M \\ \hline 0 & C - \alpha_j E_{d-a_j} \end{array}\right)$$

und wegen $B - \alpha_j E_{a_j} = \begin{pmatrix} 0 & * & * \\ & \ddots & * \\ 0 & & 0 \end{pmatrix} \in M_{a_j}(\mathbb{K})$ ist für alle $l \geq a_j$ die Potenz $(B - \alpha_j E_{a_j})^l = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$. Ferner ist

$$(C - \alpha_j E_{d-a_j})^l = \begin{pmatrix} (\alpha_1 - \alpha_j)^l & * & & \\ & \ddots & * & \\ & & (\alpha_1 - \alpha_j)^l & * \\ & & & \ddots \end{pmatrix} \in M_{d-a_j}(\mathbb{K})$$

eine rechte obere Dreiecksmatrix, die auf der Diagonale nur Einträge $(\alpha_k - \alpha_j)^l \neq 0$ mit $k \neq j$ enthält. Für $l \geq a_j$ ist also $(C - \alpha_j E_{d-a_j})^l$ als rechte obere Dreiecksmatrix mit von Null verschiedenen Einträgen auf der Diagonalen in Zeilenstufenform und damit

$$\begin{aligned} \operatorname{Rang}((A - \alpha_j E_d)^l) &= \operatorname{Rang}((A - \alpha_j E_d)^l) \\ &= \operatorname{Rang} \left( \begin{array}{c|c} 0 & * \quad\quad * \quad\quad * \\ \hline 0 & (C - \alpha_j E_{d-a_j})^l \\ & \end{array} \right) = d - a_j \end{aligned}$$

Nach der Dimensionsformel ist $\dim_{\mathbb{K}}(\operatorname{Kern}((A - \alpha_j E_d)^l)) = d - \operatorname{Rang}((A - \alpha_j E_d)^l) = a_j$ für $l \geq a_j$. Da nach Konstruktion des Hauptraums

$$\mathcal{H}(\alpha_j) = \operatorname{Kern}\left[(F - \alpha_j \mathrm{id}_V)^{\nu_j}\right] = \operatorname{Kern}\left[(F - \alpha_j \mathrm{id}_V)^l\right]$$

für $l \geq \nu_j$ und somit auch für $l \geq \max\{a_j, \nu_j\}$ ist, folgt $\dim_{\mathbb{K}}(H(\alpha_j)) = a_j$.

- Ist $v \in \operatorname{Kern}\left[(F - \alpha_j \mathrm{id}_V)^{\nu_j}\right] \cap \left[(F - \alpha_j \mathrm{id}_V)^{\nu_j}\right](V)$, so ist $(F - \alpha_j \mathrm{id}_V)^{\nu_j} v = \mathbf{0}$ und $v = (F - \alpha_j \mathrm{id}_V)^{\nu_j} w$ für ein $w \in V$. Dann ist $(F - \alpha_j \mathrm{id}_V)^{2\nu_j} w = (F - \alpha_j \mathrm{id}_V)^{\nu_j} v = \mathbf{0}$, also $w \in \operatorname{Kern}\left[(F - \alpha_j \mathrm{id}_V)^{2\nu_j}\right] = \operatorname{Kern}\left[(F - \alpha_j \mathrm{id}_V)^{\nu_j}\right]$, dh. $v = (F - \alpha_j \mathrm{id}_V)^{\nu_j} w = \mathbf{0}$. Das zeigt $\operatorname{Kern}\left[(F - \alpha_j \mathrm{id}_V)^{\nu_j}\right] \cap \left[(F - \alpha_j \mathrm{id}_V)^{\nu_j}\right](V) = \{\mathbf{0}\}$. Nach der Dimensionsformel in Satz 8.1.7 angewandt auf die lineare Abbildung $(F - \alpha_j \mathrm{id}_V)^{\nu_j}$ folgt

$$d = \dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(\operatorname{Kern}((F - \alpha_j \mathrm{id}_V)^{\nu_j})) + \dim_{\mathbb{K}}(((F - \alpha_j \mathrm{id}_V)^{\nu_j})(V)). \tag{11.3.10}$$

Mit (11.3.10) und der Dimensionsformel aus Satz 7.2.9 angewandt auf die Untervektorräume $\operatorname{Kern}\left[(F - \alpha_j \mathrm{id}_V)^{\nu_j}\right]$ und $\left[(F - \alpha_j \mathrm{id}_V)^{\nu_j}\right](V)$ folgt:

$$\begin{aligned} &\dim_{\mathbb{K}}\Big(\operatorname{Kern}((F - \alpha_j \mathrm{id}_V)^{\nu_j}) + ((F - \alpha_j \mathrm{id}_V)^{\nu_j})(V)\Big) \\ &= \dim_{\mathbb{K}}\Big(\operatorname{Kern}((F - \alpha_j \mathrm{id}_V)^{\nu_j})\Big) + \dim_{\mathbb{K}}\Big(((F - \alpha_j \mathrm{id}_V)^{\nu_j})(V)\Big) \\ &\quad - \dim_{\mathbb{K}}\Big(\operatorname{Kern}((F - \alpha_j \mathrm{id}_V)^{\nu_j}) \cap ((F - \alpha_j \mathrm{id}_V)^{\nu_j})(V)\Big) = d, \end{aligned}$$

was wegen $\operatorname{Kern}((F - \alpha_j \mathrm{id}_V)^{\nu_j}) + ((F - \alpha_j \mathrm{id}_V)^{\nu_j})(V) \subseteq V$ auch noch $V = \operatorname{Kern}((F - \alpha_j \mathrm{id}_V)^{\nu_j}) + ((F - \alpha_j \mathrm{id}_V)^{\nu_j})(V)$ und damit (11.3.9) zeigt.

- Induktion nach $m$:
  Induktionsanfang $m = 1$. Im Fall $m = 1$ ist $p_F = (\alpha_1 - X)^d$ das charakteristische Polynom von $F$, daher folgt aus $\mathcal{H}(\alpha_1) \subseteq V$ und $\dim_{\mathbb{K}}(\mathcal{H}(\alpha_1)) = a_1 = d$ die Gleichheit $\mathcal{H}(\alpha_1) = V$.
  Induktionsschritt: $m \to m + 1$:
  Es sei $p_F(X) = (\alpha_1 - X)^{a_1} \cdots (\alpha_m - X)^{a_m} (\alpha_{m+1} - X)^{a_{m+1}}$ das charakteristische Polynom von $F$. Nach (11.3.9) ist

  $$V = \operatorname{Kern}((F - \alpha_{m+1}\mathrm{id}_V)^{\nu_{m+1}}) \oplus ((F - \alpha_{m+1}\mathrm{id}_V)^{\nu_{m+1}})(V) =: \mathcal{H}(\alpha_{m+1}) \oplus W.$$

  Wie eben gezeigt ist $F(\mathcal{H}(\alpha_{m+1})) \subseteq \mathcal{H}(\alpha_{m+1})$ und ebenso ist $F(W) \subseteq W$. Wir können daher die Einschränkung $G := F|_W$ von $F$ auf $W$ betrachten, dann ist $p_G(X) = (\alpha_1 - X)^{a_1} \cdots (\alpha_m - X)^{a_m}$ das charakteristische Polynom von $G$ und nach Induktionsannahme ist dann $W = \mathcal{H}(\alpha_1) \oplus ... \oplus \mathcal{H}(\alpha_m)$ und damit

  $$V = \mathcal{H}(\alpha_1) \oplus ... \oplus \mathcal{H}(\alpha_m) \oplus \mathcal{H}(\alpha_{m+1}).$$

□

**Bemerkung 11.3.7.** Ist in der Situation von Lemma 11.3.6 für $j \in \{1, ..., m\}$ eine Basis $\mathcal{B}_j = \{v_{1,\alpha_j}, ..., v_{a_j,\alpha_j}\}$ von $\mathcal{H}(\alpha_j)$ gegeben, dann ist wegen $V = \mathcal{H}(\alpha_1) \oplus ... \oplus \mathcal{H}(\alpha_m)$

$$\mathcal{B} = \{v_{1,\alpha_1}, ..., v_{a_1,\alpha_1}, ..., v_{1,\alpha_m}, ..., v_{a_m,\alpha_m}\}$$

eine Basis von $V$ und die darstellende Matrix von $F$ bezüglich der Basis $\mathcal{B}$ ist wegen $F(\mathcal{H}(\alpha_j)) \subseteq \mathcal{H}(\alpha_j)$ immer eine Matrix in Blockdiagonalform

$$M_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} \boxed{B_1} & & \\ & \ddots & \\ & & \boxed{B_m} \end{pmatrix}$$

mit $B_j \in M_{a_j}(\mathbb{K})$ für $j = 1, ..., m$.

**Lemma 11.3.8.** *Es sei $F : V \to V$ $\mathbb{K}-$linear und $\alpha_j \in \mathbb{K}$ ein Eigenwert von $F$. Ist $\operatorname{Eig}(F, \alpha_j, l) = \operatorname{Eig}(F, \alpha_j, l-1) \oplus \operatorname{lin}\{y_1, ..., y_m\}$ mit linear unabhängigen Vektoren $y_1, ..., y_m$, dann gibt es $q \in \mathbb{N}_0$ und im Fall $q > 0$ Elemente $x_1, ..., x_q \in \operatorname{Eig}(F, \alpha_j, l - 1)$ so daß*

$$\operatorname{Eig}(F, \alpha_j, l - 1) = \operatorname{Eig}(F, \alpha_j, l - 2) \oplus \operatorname{lin}\{(F - \alpha_j)y_1, ...(F - \alpha_j)y_m, x_1, ..., x_q\}$$

*mit linear unabhängigen $(F - \alpha_j)y_1, ..., (F - \alpha_j)y_m, x_1, ..., x_q$.*

*Beweis.* Nach Voraussetzung sind $y_1, ..., y_m \in \operatorname{Eig}(F, \alpha_j, l)$, dh.

$$(F - \alpha_j)^l(y_k) = (F - \alpha_j)^{l-1}((F - \alpha_j)(y_k)) = \mathbf{0} \qquad \text{für } k = 1, ..., m$$

also $(F - \alpha_j)(y_k) \in \operatorname{Eig}(F, \alpha_j, l - 1)$ für $k = 1, ..., m$. Es seien $\lambda_1, ..., \lambda_m \in \mathbb{K}$ mit $\sum_{k=1}^{m} \lambda_k(F - \alpha_j)(y_k) \in \operatorname{Eig}(F, \alpha_j, l - 2)$, dh.

$$\mathbf{0} = (F - \alpha_j)^{l-2} \left( \sum_{k=1}^{m} \lambda_k(F - \alpha_j)(y_k) \right) = (F - \alpha_j)^{l-1} \left( \sum_{k=1}^{m} \lambda_k y_k \right)$$

also gilt

$$\sum_{k=1}^{m} \lambda_k y_k \in \mathrm{Eig}(F, \alpha_j, l-1) \cap \mathrm{lin}\{y_1, ..., y_m\}. \tag{11.3.11}$$

Weil es nach Voraussetzung die direkte Summenzerlegung

$$\mathrm{Eig}(F, \alpha_j, l) = \mathrm{Eig}(F, \alpha_j, l-1) \oplus \mathrm{lin}\{y_1, ..., y_m\}$$

gibt, folgt insbesondere $\mathrm{Eig}(F, \alpha_j, l-1) \cap \mathrm{lin}\{y_1, ..., y_m\} = \{\mathbf{0}\}$ und deshalb folgt

$$\sum_{k=1}^{m} \lambda_k y_k = \mathbf{0}$$

aus (11.3.11). Da $y_1, ..., y_m$ linear unabhängig sind, folgt $\lambda_1 = ... = \lambda_m = 0$. Dies zeigt:

- $\mathrm{Eig}(F, \alpha_j, l-2) \cap \mathrm{lin}\{(F-\alpha_j)(y_1), ..., (F-\alpha_j)(y_m)\} = \{\mathbf{0}\}$
- $(F-\alpha_j)(y_1), ..., (F-\alpha_j)(y_m)$ sind linear unabhängig.

Daher ist $W := \mathrm{Eig}(F, \alpha, l-2) \oplus \mathrm{lin}\{(F-\alpha_j)(y_1), ...., (F-\alpha_j)(y_m)\}$ wohldefiniert und $W \subseteq \mathrm{Eig}(F, \alpha_j, l-1)$. Damit läßt sich jede Basis $\mathcal{B}_{l-2} = (v_1, ..., v_p)$ von $\mathrm{Eig}(F, \alpha_j, l-2)$ zu einer Basis $\mathcal{B}_W = (v_1, ..., v_p, (F-\alpha_j)(y_1), ...., (F-\alpha_j)(y_m))$ von $W$ und – falls $W \neq \mathrm{Eig}(F, \alpha_j, l-1)$ – weiter zu einer Basis

$$\mathcal{B}_{l-1} = (v_1, ..., v_p, (F-\alpha_j)(y_1), ...., (F-\alpha_j)(y_m), x_1, ..., x_q)$$

von $\mathrm{Eig}(F, \alpha_j, l-1)$ ergänzen. $\square$

**Satz 11.3.9.** *Es sei $V$ ein $d$-dimensionaler $\mathbb{K}$-Vektorraum und $F : V \to V$ sei $\mathbb{K}$-linear und $A = M_{\mathcal{A}}^{\mathcal{A}}(F) \in M_d(\mathbb{K})$ die darstellende Matrix von $F$ in einer Basis $\mathcal{A}$ von $V$. Das charakteristische Polynom $p_F$*

$$p_F(X) = \det(A - XE_d) = (\alpha_1 - X)^{a_1} \cdots (\alpha_m - X)^{a_m} \tag{11.3.12}$$

*zerfalle in ein Produkt von Linearfaktoren. Wir setzen in (11.3.12) für $F$ paarweise verschiedene Eigenwerte $\alpha_1, ..., \alpha_m \in \mathbb{K}$ und $a_1, ..., a_m \in \mathbb{N}$ mit $a_1 + ... + a_m = d$ voraus. Dann gilt:*

*a) $F$ ist genau dann diagonalisierbar wenn*

$$a_j = \dim(\mathrm{Kern}(A - \alpha_j E_d)) =: g_j \quad \textit{für} \quad j = 1, ..., m \tag{11.3.13}$$

**“algebraische Vielfachheit = geometrische Vielfachheit”** *ist. In diesem Fall gibt es eine Basis $\mathcal{B}$ von $V$, in der die darstellende Matrix $M_{\mathcal{B}}^{\mathcal{B}}(F)$ eine Diagonalmatrix der Form*

$$M_{\mathcal{B}}^{\mathcal{B}}(F) = T^{-1}AT = \begin{pmatrix} \boxed{\alpha_1 E_{a_1}} & & & \\ & \boxed{\alpha_2 E_{a_2}} & & \\ & & \ddots & \\ & & & \boxed{\alpha_m E_{a_m}} \end{pmatrix}$$

*ist.*

b) *Es gibt zu $F$ immer eine Basis $\mathcal{B}$ von $V$, so daß $M^{\mathcal{B}}_{\mathcal{B}}(F)$ Jordan Normalform besitzt. Die Basiselemente lassen sich so anordnen, daß die Jordan Normalform*

$$J = M^{\mathcal{B}}_{\mathcal{B}}(F) = \begin{pmatrix} \boxed{J_1} & & \\ & \ddots & \\ & & \boxed{J_m} \end{pmatrix} = T^{-1}AT \tag{11.3.14}$$

*sich als Blockdiagonalmatrix mit Jordanblöcken $J_1 \in M_{a_1}(\mathbb{K})$ zum Eigenwert $\alpha_1$, $J_m \in M_{a_m}(\mathbb{K})$ zum Eigenwert $\alpha_m$ schreibt, wobei sich jeder der Jordanblöcke $J_k$ in $g_k$ Jordankastl ($k = 1, ..., m$) aufteilt.*

*In beiden Fällen [1] sind die Transformationsmatrizen $T^{-1} = M^{\mathcal{A}}_{\mathcal{B}}(\mathrm{id}_V)$ bzw. $T = (M^{\mathcal{A}}_{\mathcal{B}}(\mathrm{id}_V))^{-1} = M^{\mathcal{B}}_{\mathcal{A}}(\mathrm{id}_V)$.*

*Beweis von Satz 11.3.9 und Konstruktionsrezept der Jordan Normalform.*
Für $j \in \{1, ..., m\}$ sei $\mathcal{C}_j = \{w_{1,\alpha_j}, ..., w_{a_j,\alpha_j}\}$ eine Basis von $\mathcal{H}(\alpha_j)$, dann ist die darstellende Matrix von $F$ bezüglich der Basis

$$\mathcal{C} = \{w_{1,\alpha_1}, ..., w_{a_1,\alpha_1}, ..., w_{1,\alpha_m}, ..., w_{a_m,\alpha_m}\}$$

wegen $F(\mathcal{H}(\alpha_j)) \subseteq \mathcal{H}(\alpha_j)$ eine Matrix in Blockdiagonalform

$$M^{\mathcal{C}}_{\mathcal{C}}(F) = \begin{pmatrix} \boxed{B_1} & & \\ & \ddots & \\ & & \boxed{B_m} \end{pmatrix}$$

mit $B_j \in M_{a_j}(\mathbb{K})$ für $j = 1, ..., m$. Durch spezielle Wahl der Basen $\mathcal{C}_j$ von $\mathcal{H}(\alpha_j)$ läßt sich Diagonalform bzw. Jordan Normalform erreichen:

a) Gilt $a_j = g_j$, also $\dim_{\mathbb{K}}(\mathcal{H}(\alpha_j)) = \dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j))$, so folgt $\mathrm{Eig}(F, \alpha_j) = \mathcal{H}(\alpha_j)$ aus (11.3.6), damit ist $F|_{\mathcal{H}(\alpha_j)}$ diagonalisierbar und die darstellende Matrix von $F|_{\mathcal{H}(\alpha_j)}$ bezüglich einer beliebigen Basis $\mathcal{B}^{(j)}$ von $\mathcal{H}(\alpha_j)$ ist die Diagonalmatrix

$$M^{\mathcal{B}^{(j)}}_{\mathcal{B}^{(j)}}(F|_{\mathcal{H}(\alpha_j)}) = \begin{pmatrix} \alpha_j & & \\ & \ddots & \\ & & \alpha_j \end{pmatrix} \in M_{a_j}(\mathbb{K}).$$

b) Gilt $g_j < a_j$, so gibt es keine Basis von $\mathcal{H}(\alpha_j)$ bestehend aus Eigenvektoren, also ist $F|_{\mathcal{H}(\alpha_j)}$ nicht diagonalisierbar. Nach Lemma 11.3.4 gibt es ein $\nu_j > 1$ mit

$$\{0\} \subsetneq \mathrm{Eig}(F, \alpha_j) \subsetneq \cdots \subsetneq \mathrm{Eig}(F, \alpha_j, \nu_j - 1) \subsetneq \mathrm{Eig}(F, \alpha_j, \nu_j) = \mathcal{H}(\alpha_j).$$

Ergänze eine Basis $(\xi_1, ..., \xi_N)$ von $\mathrm{Eig}(F, \alpha_j, \nu_j - 1)$ zu einer Basis $(\xi_1, ..., \xi_N, \xi_{N+1}, ..., \xi_{a_j})$ von $\mathrm{Eig}(F, \alpha_j, \nu_j) = \mathcal{H}(\alpha_j)$, dann ist $v_{1,\nu_j} := \xi_{N+1} \in \mathrm{Eig}(F, \alpha_j, \nu_j) \backslash \mathrm{Eig}(F, \alpha_j, \nu_j - 1)$

[1] Mit dieser Konvention enthält $T = M^{\mathcal{B}}_{\mathcal{A}}(\mathrm{id}_V)$ als Spaltenvektoren die Koordinatenvektoren der Basiselemente von $\mathcal{B}$ ausgedrückt durch die Basiselemente von $\mathcal{A}$.

ein Hauptvektor von maximaler, also $\nu_j$-ter Stufe und berechne sukzessive

$$\begin{aligned} v_{1,\nu_j-1} &:= (F-\alpha_j \mathrm{id}_V)v_{1,\nu_j} \in \mathrm{Eig}(F,\alpha_j,\nu_j-1)\backslash\mathrm{Eig}(F,\alpha_j,\nu_j-2) && (11.3.15)\\ \vdots \quad & \quad \vdots \\ v_{1,2} &:= (F-\alpha_j \mathrm{id}_V)v_{1,3} \in \mathrm{Eig}(F,\alpha_j,2)\backslash\mathrm{Eig}(F,\alpha_j) && (11.3.16)\\ v_{1,1} &:= (F-\alpha_j \mathrm{id}_V)v_{1,2} \in \mathrm{Eig}(F,\alpha_j)\backslash\{\mathbf{0}\}. && (11.3.17) \end{aligned}$$

Dann sind $v_{1,1},...,v_{1,\nu_j}$ linear unabhängig: Ist $\lambda_1 v_{1,1}+...+\lambda_{\nu_j}v_{1,\nu_j}=\mathbf{0}$, so folgt durch Anwenden von $(F-\alpha_j\mathrm{id}_V)$:

$$\begin{aligned} \mathbf{0} &= (F-\alpha_j\mathrm{id}_V)(\mathbf{0})\\ &= \lambda_1(F-\alpha_j\mathrm{id}_V)(v_{1,1})+...+\lambda_{\nu_j}(F-\alpha_j\mathrm{id}_V)(v_{1,\nu_j})\\ &= \lambda_2(F-\alpha_j\mathrm{id}_V)(v_{1,2})+...+\lambda_{\nu_j}(F-\alpha_j\mathrm{id}_V)(v_{1,\nu_j}) \end{aligned}$$

denn $(F-\alpha_j\mathrm{id}_V)(v_{1,1})=\mathbf{0}$. Durch wiederholte Anwendung von $F-\alpha_j\mathrm{id}_V$ werden sukzessive $(F-\alpha_j\mathrm{id}_V)^2(v_{1,2})=\mathbf{0}$, ..., $(F-\alpha_j\mathrm{id}_V)^{\nu_j-1}(v_{1,\nu_j-1})=\mathbf{0}$ und wir erhalten damit

$$\begin{aligned} \mathbf{0} &= \lambda_3(F-\alpha_j\mathrm{id}_V)^2(v_{1,3})+...+\lambda_{\nu_j}(F-\alpha_j\mathrm{id}_V)^2(v_{1,\nu_j})\\ \vdots\ & \quad \vdots\\ \mathbf{0} &= \lambda_{\nu_j-1}(F-\alpha_j\mathrm{id}_V)^{\nu_j-2}(v_{1,\nu_j-1})+\lambda_{\nu_j}(F-\alpha_j\mathrm{id}_V)^{\nu_j-2}(v_{1,\nu_j})\\ \mathbf{0} &= \lambda_{\nu_j}(F-\alpha_j\mathrm{id}_V)^{\nu_j-1}(v_{1,\nu_j}). \end{aligned}$$

Wegen $v_{1,k}\in\mathrm{Eig}(F,\alpha_j,k)\backslash\mathrm{Eig}(F,\alpha_j,k-1)$ folgt zunächst $(F-\alpha_j\mathrm{id}_V)^{\nu_j-1}(v_{1,\nu_j})\neq\mathbf{0}$, also $\lambda_{\nu_j}=0$ und dann sukzessive $\lambda_{\nu_j-1}=0$,..., $\lambda_2=0$, $\lambda_1=0$.

Aus $F(v_{1,1})=\alpha_j v_{1,1}$ und durch Auflösen von (11.3.17) folgt dann

$$F(v_{1,2})=v_{1,1}+\alpha_j v_{1,2}$$

und sukzessive aus (11.3.16) bis (11.3.15):

$$F(v_{1,3})=v_{1,2}+\alpha_j v_{1,3},...,F(v_{1,\nu_j})=\alpha_j v_{1,\nu_j}+v_{1,\nu_j-1}$$

Für $V_1:=\mathrm{lin}\{v_{1,1},...,v_{1,\nu_j}\}$ gilt daher $F(V_1)\subseteq V_1$ und $F|_{V_1}$ hat bezüglich der Basis $\mathcal{B}_1:=(v_{1,1},...,v_{1,\nu_j})$ von $V_1$ das Jordankastl

$$M^{\mathcal{B}_1}_{\mathcal{B}_1}(F|_{V_1})=\begin{pmatrix} \alpha_j & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1\\ & & & \alpha_j \end{pmatrix}\in M_{\nu_j}(\mathbb{K})$$

als darstellende Matrix. Ist nun $V_1\neq\mathcal{H}(\alpha_j)$, so gibt es bei den Inklusionen

$$\{0\}\subsetneq\mathrm{Eig}(F,\alpha_j)\subsetneq\cdots\subsetneq\mathrm{Eig}(F,\alpha_j,\nu_j-1)\subsetneq\mathrm{Eig}(F,\alpha_j,\nu_j)=\mathcal{H}(\alpha_j).$$

mindestens einen Index $k \in \{1, 2, ..., \nu_j\}$ mit

$$2 + \dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j, k-1)) \leq \dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j, k)) \tag{11.3.18}$$

für $k \geq 2$, bzw.

$$\dim(\mathrm{Eig}(F, \alpha_j)) \geq 2 \tag{11.3.19}$$

im Fall $k = 1$. Wähle nun den maximalen Index

$$\mu_j := \max\{k \in \{1, ..., \nu_j\} : k \text{ erfüllt (11.3.18) oder (11.3.19) }\},$$

dann ist

$$\dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j, l)) \geq \dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j, l-1)) + 2$$

für $l \leq \mu_j$ laut Lemma 11.3.8, also

$$\mathrm{lin}\{v_{1,1}, ..., v_{1,\mu_j}\} \subsetneq \mathrm{Eig}(F, \alpha_j, \mu_j)$$

und

$$\mathrm{Eig}(F, \alpha_j, \mu_j - 1) \oplus \mathrm{lin}\{v_{1,\mu_j}\} \subsetneq \mathrm{Eig}(F, \alpha_j, \mu_j).$$

Ergänze eine Basis $(v_{1,1}, ..., v_{1,\mu_j}, \xi_1, ..., \xi_{N_1})$ von $\mathrm{Eig}(F, \alpha_j, \mu_j - 1) \oplus \mathrm{lin}\{v_{1,\mu_j}\}$ zu einer Basis $(v_{1,1}, ..., v_{1,\mu_j}, \xi_1, ..., \xi_{N_1}, \zeta_1, ..., \zeta_{N_2})$ von $\mathrm{Eig}(F, \alpha_j, \mu_j)$ und wähle

$$v_{2,\mu_j} := \zeta_1 \in \mathrm{Eig}(F, \alpha_j, \mu_j) \backslash \mathrm{Eig}(F, \alpha_j, \mu_j - 1)$$

Da $v_{1,\mu_j}, v_{2,\mu_j} \in \mathrm{Eig}(F, \alpha_j, \mu_j) \backslash \mathrm{Eig}(F, \alpha_j, \mu_j - 1)$ linear unabhängig sind, gilt dies nach Lemma 11.3.8 auch für

$$\begin{array}{rcl} v_{1,\mu_j-1} := (F - \alpha_j)v_{1,\mu_j}, v_{2,\mu_j-1} := (F - \alpha_j)v_{2,\mu_j} & \in & \mathrm{Eig}(F, \alpha_j, \mu_j - 1) \backslash \mathrm{Eig}(F, \alpha_j, \mu_j - 2) \\ & \vdots & \\ v_{1,1} := (F - \alpha_j)v_{1,2}, v_{2,1} := (F - \alpha_j)v_{2,2} & \in & \mathrm{Eig}(F, \alpha_j) \backslash \{\mathbf{0}\} \end{array}$$

$v_{1,1}, ..., v_{1,\nu_j}, v_{2,1}, ..., v_{2,\mu_j}$ sind linear unabhängig, denn für $\lambda_1, ..., \lambda_{\nu_j}, \beta_1, ..., \beta_{\mu_j} \in \mathbb{K}$ mit

$$\lambda_1 v_{1,1} + ... + \lambda_{\nu_j} v_{1,\nu_j} + \beta_1 v_{2,1} + ... + \beta_{\mu_j} v_{2,\mu_j} = \mathbf{0}$$

folgt durch wiederholtes Anwenden von $F - \alpha_j \mathrm{id}_V$:

$$\begin{aligned} 0 &= \lambda_2(F - \alpha_j \mathrm{id}_V)v_{1,2} + ... + \lambda_{\nu_j}(F - \alpha_j \mathrm{id}_V)v_{1,\nu_j} \\ &\quad + \beta_2(F - \alpha_j \mathrm{id}_V)v_{2,2} + ... + \beta_{\mu_j}(F - \alpha_j \mathrm{id}_V)v_{2,\mu_j} && (11.3.20) \\ &\vdots \\ 0 &= \lambda_{\mu_j}(F - \alpha_j \mathrm{id}_V)^{\mu_j - 1}v_{1,\mu_j} + ... + \lambda_{\nu_j}(F - \alpha_j \mathrm{id}_V)^{\mu_j - 1}v_{1,\nu_j} \\ &\quad + \beta_{\mu_j}(F - \alpha_j \mathrm{id}_V)^{\mu_j - 1}v_{2,\mu_j} && (11.3.21) \\ 0 &= \lambda_{\mu_j+1}(F - \alpha_j \mathrm{id}_V)^{\mu_j}v_{1,\mu_j+1} + ... + \lambda_{\nu_j}(F - \alpha_j \mathrm{id}_V)^{\mu_j}v_{1,\nu_j} && (11.3.22) \\ &\vdots \\ 0 &= \lambda_{\nu_j-1}(F - \alpha_j \mathrm{id}_V)^{\nu_j - 2}v_{1,\nu_j-1} + \lambda_{\nu_j}(F - \alpha_j \mathrm{id}_V)^{\nu_j - 2}v_{1,\nu_j} && (11.3.23) \\ 0 &= \lambda_{\nu_j}(F - \alpha_j \mathrm{id}_V)^{\nu_j - 1}v_{1,\nu_j} && (11.3.24) \end{aligned}$$

Wegen $(F-\alpha_j \mathrm{id}_V)^{k-1} v_{1,k} \neq \mathbf{0}$ ergibt sich $\lambda_{\nu_j} = 0 = ... = \lambda_{\mu_j+1}$ aus (11.3.24) bis (11.3.22), was dann (11.3.21) zu

$$0 = \lambda_{\mu_j}(F-\alpha_j \mathrm{id}_V)^{\mu_j-1} v_{1,\mu_j} + \beta_{\mu_j}(F-\alpha_j \mathrm{id}_V)^{\mu_j-1} v_{2,\mu_j} = \lambda_{\mu_j} v_{1,1} + \beta_{\mu_j} v_{2,1}$$

vereinfacht. Da $v_{1,1}$ und $v_{2,1}$ linear unabhängig sind, ist $\lambda_{\mu_j} = \beta_{\mu_j} = 0$. Einsetzen in die Gleichungen bis (11.3.20) ergibt noch $\lambda_2 = \beta_2 = 0 = ... = \lambda_{\mu_j-1} = \beta_{\mu_j-1}$.

Damit definiert $V_2 := \mathrm{lin}\{v_{2,1}, ..., v_{2,\mu_j}\} \subseteq \mathcal{H}(\alpha_j)$ einen Unterraum, so daß $F(V_2) \subseteq V_2$ und $F|_{V_2}$ bezüglich der Basis $\mathcal{B}_2 := (v_{2,1}, ..., v_{2,\mu_j})$ von $V_2$ das Jordankastl

$$M^{\mathcal{B}_2}_{\mathcal{B}_2}(F|_{V_2}) = \begin{pmatrix} \alpha_j & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \alpha_j \end{pmatrix} \in M_{\mu_j}(\mathbb{K})$$

als darstellende Matrix besitzt. Da $v_{1,1}, ..., v_{1,\nu_j}, v_{2,1}, ..., v_{2,\mu_j}$ linear unabhängig sind, ist die direkte Summe $V_1 \oplus V_2$ definiert. Ist $V_1 \oplus V_2 = \mathcal{H}(\alpha_j)$, so ist man fertig.

Ansonsten sucht man analog zu (11.3.18) und (11.3.19) den maximalen Index

$$\delta_j := \max\Big\{k \in \{1, ..., \mu_j\} : 3 + \dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j, k-1)) \leq \dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j, k))\Big\}$$

ergänze eine Basis $(v_{1,1}, ..., v_{1,\delta_j}, v_{2,1}, ..., v_{2,\delta_j}, \xi_1, ..., \xi_{N_3})$ von $\mathrm{Eig}(F, \alpha_j, \delta_j - 1) \oplus \mathrm{lin}\{v_{1,\delta_j}, v_{2,\delta_j}\}$ zu einer Basis $(v_{1,1}, ..., v_{1,\delta_j}, v_{2,1}, ..., v_{2,\delta_j}, \xi_1, ..., \xi_{N_3}, \zeta_1, ..., \zeta_{N_4})$ von $\mathrm{Eig}(F, \alpha_j, \delta_j)$, wähle $v_{3,\delta_j} := \zeta_1$ und wiederhole obiges Verfahren.

Da $\mathcal{H}(\alpha_j)$ endlichdimensional ist, bricht dieses Verfahren nach endlich vielen Wiederholungen ab. Ist nun $\mathcal{H}(\alpha_j) = V_1 \oplus ... \oplus V_k$, dann ist $\mathrm{Eig}(F, \alpha_j) = \mathrm{lin}\{v_{1,1}, v_{2,1}, ..., v_{k,1}\}$, also $k = g_j$ die geometrische Vielfachheit des Eigenwerts $\alpha_j$, also

$$\mathcal{H}(\alpha_j) = V_1 \oplus ... \oplus V_{g_j}.$$

Sind $\mathcal{B}_1 = (v_{1,1}, ..., v_{1,\nu_j}), ..., \mathcal{B}_{g_j} = (v_{g_j,1}, ..., v_{g_j,\kappa_j})$ die nach obiger Konstruktion bestimmten Basen der Unterräume $V_1, ..., V_{g_j}$, dann ist

$$\mathcal{B}^{(j)} = (v_{1,1}, ..., v_{1,\nu_j}, ..., v_{g_j,1}, ..., v_{g_j,\kappa_j})$$

eine Basis von $\mathcal{H}(\alpha_j)$ in der $F|_{\mathcal{H}(\alpha_j)}$ Jordan Normalform besitzt. Die darstellende Matrix

$$M^{\mathcal{B}^{(j)}}_{\mathcal{B}^{(j)}}(F|_{\mathcal{H}(\alpha_j)}) = \begin{pmatrix} \boxed{\alpha_j E_{\nu_j}} & & & \\ & \boxed{\alpha_j E_{\mu_j}} & & \\ & & \ddots & \\ & & & \boxed{\alpha_j E_{\kappa_j}} \end{pmatrix}$$

besteht dann aus genau $g_j$ Jordankastl zum Eigenwert $\alpha_j$ der Größen $\nu_j, \mu_j, ..., \kappa_j$.

Da $V = \mathcal{H}(\alpha_1) \oplus ... \oplus \mathcal{H}(\alpha_m)$ und $F(\mathcal{H}(\alpha_j)) \subseteq \mathcal{H}(\alpha_j)$ für $j = 1, ..., m$ nach Lemma 11.3.6 gilt, ist $F$ genau dann diagonalisierbar, wenn jedes $F|_{\mathcal{H}(\alpha_j)}$, $j = 1, ..., m$ diagonalisierbar ist, dh. wenn für jeden der Eigenwerte $\alpha_1, ..., \alpha_m$ geometrische und algebraische Vielfachheit übereinstimmen. Anwendung von Satz 8.3.11 gibt

$$M_{\mathcal{B}}^{\mathcal{B}}(F) = M_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id}_V) M_{\mathcal{A}}^{\mathcal{A}}(F)(M_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id}_V))^{-1} = T^{-1}AT. \qquad \square$$

**Bemerkung 11.3.10.**

a) Wir bemerken an dieser Stelle, daß (11.3.12) für $\mathbb{K} = \mathbb{C}$ nach dem Fundamentalsatz der Algebra immer erfüllt werden kann.

b) Gilt (11.3.12), so lassen sich die Basiselemente von $\mathcal{B}$ so anordnen, daß die Jordan Normalform

$$J = M_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} \boxed{J_1} & & \\ & \ddots & \\ & & \boxed{J_m} \end{pmatrix}$$

aus $m$ **Jordanblöcken** $J_1 \in M_{a_1}(\mathbb{K}), ..., J_m \in M_{a_m}(\mathbb{K})$ zu den verschiedenen Eigenwerten besteht. Die algebraische Vielfachheit $a_j$ des Eigenwerts $\alpha_j$ – abzulesen aus der Faktorisierung (11.3.12) des charakteristischen Polynoms – gibt gerade die Größe des Jordanblocks $J_j$ an. Jeder Jordanblock $J_j$ zum Eigenwert $\alpha_j$ ist aus $g_j := \dim_{\mathbb{K}}(\mathrm{Eig}(F, \alpha_j))$ verschiedenen Jordankastl $\begin{pmatrix} \alpha_j & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \alpha_j \end{pmatrix}$ aufgebaut. Es gibt daher immer $\#_1, ..., \#_{a_j-1} \in \{0, 1\}$, so daß $J_j = \begin{pmatrix} \alpha_j & \#_1 & & \\ & \ddots & \ddots & \\ & & \ddots & \#_{a_j-1} \\ & & & \alpha_j \end{pmatrix}$ ist.

**Beispiel 11.3.11.** Für die Matrix

$$J = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

in Jordan Normalform liest man sofort – aus

$$J + E_9 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

,

$$(J + E_9)^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

bzw. $(J + E_9)^3 = 0$ – ab:

$$\begin{aligned} \mathrm{Eig}(J, -1) &= \mathrm{lin}\{\underline{e}_1, \underline{e}_4, \underline{e}_7, \underline{e}_9\} \subsetneq \mathrm{Eig}(J, -1, 2) = \mathrm{lin}\{\underline{e}_1, \underline{e}_2, \underline{e}_4, \underline{e}_5, \underline{e}_7, \underline{e}_8, \underline{e}_9\} \\ &\subsetneq \mathrm{Eig}(J, -1, 3) = \mathrm{lin}\{\underline{e}_1, \underline{e}_2, \underline{e}_3, \underline{e}_4, \underline{e}_5, \underline{e}_6, \underline{e}_7, \underline{e}_8, \underline{e}_9\} = \mathcal{H}(-1) \end{aligned}$$

Bei der Durchführung des Verfahrens von Beweis des Satzes 11.3.9 läßt sich also im ersten Schritt $v_{1,3} := \underline{e}_3 \in H(-1)\backslash \mathrm{Eig}(J, -1, 2)$ wählen, dann sind $v_{1,2} := (J + E_9)v_{1,3} = \underline{e}_2$ und $v_{1,1} := (J + E_9)v_{1,2} = \underline{e}_1$ und $V_1 := \mathrm{lin}\{v_{1,1}, v_{1,2}, v_{1,3}\} = \mathrm{lin}\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ hier die im Beweis konstruierten Vektoren. Wegen $V_1 \neq \mathcal{H}(-1)$ muß das Verfahren wiederholt werden. Da $\dim(\mathrm{Eig}(J, -1, 2)) + 2 = \dim(\mathrm{Eig}(J, -1, 3))$ ist $\mu = 3$ und

$$v_{2,3} := \underline{e}_6 \in \mathrm{Eig}(J, -1, 3)\backslash \mathrm{Eig}(J, -1, 2)$$

ein Basisvektor der bei Ergänzung einer Basis $(v_{1,1}, v_{1,2}, v_{1,3}, \xi_1, ..., \xi_5)$ von $\mathrm{Eig}(J, -1, 2) \oplus \mathrm{lin}\{v_{1,3}\}$ zu einer Basis von $\mathrm{Eig}(J, -1, 3)$ im verallgemeinerten Eigenraum höchster verbleibender Stufe, also in $\mathrm{Eig}(J, -1, 3)$ liegt. Wiederholung dieser Konstruktion mit $v_{2,3}$ ergibt $v_{2,2} := (J+E_9)v_{2,3} = \underline{e}_5$ und $v_{2,1} := (J+E_9)v_{2,2} = \underline{e}_4$ und damit $V_2 := \mathrm{lin}\{\underline{e}_4, \underline{e}_5, \underline{e}_6\}$. Nun ist $V_1 \oplus V_2 \neq \mathcal{H}(-1)$, also muß das Verfahren weiter wiederholt werden. $v_{3,2} := \underline{e}_8$ ist ein möglicher verallgemeinerter Eigenvektor der höchsten verbleibenden Stufe – hier $\kappa = 2$ – der beim Ergänzen der Basis $(v_{1,1}, v_{1,2}, v_{2,1}, v_{2,2}, \underline{e}_7, \underline{e}_9)$ von $\mathrm{Eig}(J, -1) \oplus \mathrm{lin}\{v_{1,2}, v_{2,2}\}$ zu einer Basis von $\mathrm{Eig}(J, -1, 2)$ dazukommt. Mit dem wird das Verfahren erneut gestartet: $v_{3,1} := (J + E_9)v_{3,2} = \underline{e}_7$ und damit wird $V_3 := \mathrm{lin}\{\underline{e}_7, \underline{e}_8\}$. Da $V_1 \oplus V_2 \oplus V_3 \neq \mathcal{H}(-1)$ ist, wird das Verfahren wiederholt. $\underline{e}_9 \in \mathrm{Eig}(J, -1)$ ist ein (verallgemeinerter) Eigenvektor der höchsten verbleibenden Stufe – hier 1 – der noch nicht in $V_1 \oplus V_2 \oplus V_3$ liegt. Wegen $\mathcal{H}(-1) = V_1 \oplus V_2 \oplus V_3 \oplus \mathrm{lin}\{\underline{e}_9\}$ endet hier das Verfahren.

**Beispiel 11.3.12.** Zu $A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ -2 & -3 & 0 \end{pmatrix}$ ist das charakteristische Polynom $\det(A - \mu E_3) = \mu^2(2-\mu) + 6 - 4(2-\mu) - 3\mu = (2-\mu)(\mu-1)(\mu+1)$, damit sind alle Eigenwerte einfach, $A$ ist somit diagonalisierbar, durch elementare Zeilenumformungen wird

$$A + E_3 = \begin{pmatrix} 1 & 0 & -2 \\ 1 & 3 & 1 \\ -2 & -3 & 1 \end{pmatrix} \overset{(II)-(I),(III)+2(I)}{\longrightarrow} \begin{pmatrix} 1 & 0 & -2 \\ 0 & 3 & 3 \\ 0 & -3 & -3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

und damit ist $\mathrm{Kern}(A+E) = \mathrm{lin}\left\{\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}\right\}$ der Eigenraum von $A$ zum Eigenwert $-1$.

Analog erhalten wir $\mathrm{Kern}(A-E) = \mathrm{lin}\left\{\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}\right\}$ und $\mathrm{Kern}(A-2E) = \mathrm{lin}\left\{\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}$.

Mit $\begin{pmatrix} -1 & -2 & 2 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & -2 & 0 \\ \frac{1}{2} & \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ ergibt sich nun

$$\begin{pmatrix} -1 & -2 & 0 \\ \frac{1}{2} & \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} A \begin{pmatrix} -1 & -2 & 2 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

**Beispiel 11.3.13.** Zur Berechnung der Jordan Normalform:

$$A = \begin{pmatrix} 0 & -1 & -2 \\ -1 & 0 & 2 \\ 1 & -1 & -3 \end{pmatrix}$$

hat das charakteristische Polynom $-(x+1)^3$, also den einen Eigenwert $-1$ mit algebraischer Vielfachheit $a = 3$ und mit geometrischer Vielfachheit $g = \dim(\mathrm{Eig}(A,-1)) = 2$ und $\mathrm{Eig}(A,-1) = \mathrm{lin}\{\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\}$. Mit $(A+E)^2 = 0$ folgt $\mathcal{H}(-1) = \mathrm{Eig}(A,-1,2) = \mathbb{R}^3$, also ist wegen $\det\begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = 2 \neq 0$ der Vektor $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in \mathcal{H}(-1)\backslash\mathrm{Eig}(A,-1)$ ein Hauptvektor maximaler, also 2. Stufe. $(A+E)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ -2 \end{pmatrix}$. Somit ist $V_1 = \mathrm{lin}\left\{\begin{pmatrix} -2 \\ 2 \\ -2 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ der erste in der Konstruktion gebildete Unterraum. Da $\det\begin{pmatrix} -2 & 0 & 1 \\ 2 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix} = 4 \neq 0$ ist, wird der Raum $W_2$ von $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ aufgespannt und daher ist

$\left\{ \begin{pmatrix} -2 \\ 2 \\ -2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$ eine mögliche Basis von $\mathbb{R}^3$, in der die darstellende Matrix der linearen Abbildung $F_A$ in Jordan Normalform vorliegt:

$$J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = M_{\mathcal{B}}^{\mathcal{B}}(F_A)$$

$A$ ist also ähnlich zu $J$, dh. $A = TJT^{-1}$ mit $T, T^{-1} \in GL(3, \mathbb{R})$ und aus obiger Rechnung erhält man die „Transformationsmatrizen“ $T$ und $T^{-1}$ mit dem Satz 8.3.11:

$$A = M_{\mathcal{K}}^{\mathcal{K}}(F_A) = M_{\mathcal{K}}^{\mathcal{B}}(\mathrm{id}_{\mathbb{R}^3}) M_{\mathcal{B}}^{\mathcal{B}}(F_A) [M_{\mathcal{K}}^{\mathcal{B}}(\mathrm{id}_{\mathbb{R}^3})]^{-1},$$

woraus sich nach Definition der darstellenden Matrix und obiger Rechnung $T = M_{\mathcal{K}}^{\mathcal{B}}(\mathrm{id}_{\mathbb{R}^3}) = \begin{pmatrix} -2 & 0 & 1 \\ 2 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}$ ergibt. Nach Matrixinvertieren ergibt sich $T^{-1} = \frac{1}{4} \begin{pmatrix} -1 & 1 & 0 \\ -2 & 2 & 4 \\ 2 & 2 & 0 \end{pmatrix}$ und so $A = TJT^{-1}$.

**Bemerkung 11.3.14.** Hat $A \in M_d(\mathbb{R})$ ein charakteristisches Polynom [2]

$$p_A(X) = (\beta_1 - X)^{b_1} \cdots (\beta_k - X)^{b_k} (\gamma_1 - X)^{c_1} \cdots (\gamma_l - X)^{c_l} (\overline{\gamma_1} - X)^{c_1} \cdots (\overline{\gamma_l} - X)^{c_l} \quad (11.3.25)$$

mit paarweise verschiedenen $\beta_1, ..., \beta_k \in \mathbb{R}$ und paarweise verschiedenen $\gamma_1, ... \gamma_l \in \mathbb{C} \setminus \mathbb{R}$, $\mathrm{Im}\,(\gamma_1), ..., \mathrm{Im}\,(\gamma_l) > 0$, $b_1, ..., b_k, c_1, ..., c_l \in \mathbb{N}$, $l \geq 1$ und $b_1 + ... + b_k + 2(c_1 + ... + c_l) = d$, so gibt es keine Jordan Normalform für die $\mathbb{R}-$lineare Abbildung $\begin{array}{rccc} F_A : & \mathbb{R}^d & \to & \mathbb{R}^d \\ & \underline{x} & \mapsto & A\underline{x} \end{array}$. Faßt man $A \in M_d(\mathbb{R})$ als darstellende Matrix der $\mathbb{C}-$linearen Abbildung $\begin{array}{rccc} G_A : & \mathbb{C}^d & \to & \mathbb{C}^d \\ & \underline{y} & \mapsto & A\underline{y} \end{array}$ auf, so gibt es eine Jordan Normalform $J \in M_d(\mathbb{C})$ zu $G_A$. Sind nun

$$J_j(\gamma) = \begin{pmatrix} \gamma & 1 & 0 & \cdots & 0 \\ 0 & \gamma & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \gamma \end{pmatrix} \quad , \quad J_j(\overline{\gamma}) = \begin{pmatrix} \overline{\gamma} & 1 & 0 & \cdots & 0 \\ 0 & \overline{\gamma} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \overline{\gamma} \end{pmatrix} \in M_j(\mathbb{K})$$

zwei Jordankastl aus $J$ (mit $\gamma \in \mathbb{C}$), dann erhält man unter Verwendung von

$$S := \frac{1}{\sqrt{2}} \begin{pmatrix} E_j & E_j \\ -iE_j & iE_j \end{pmatrix} \quad \text{und} \quad S^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} E_j & iE_j \\ E_j & -iE_j \end{pmatrix} \quad (11.3.26)$$

[2] sobald $p_A$ nicht nur reelle Nullstellen hat, ist das die allgemeine Form von $p_A$, denn mit jeder Nullstelle $\gamma \in \mathbb{C} \setminus \mathbb{R}$ von $p_A \in \mathbb{C}[X]$ für $A \in M_d(\mathbb{R})$ ist dann auch die komplex konjugierte Zahl $\overline{\gamma}$ eine Nullstelle von $p_A$ mit derselben Nullstellenordnung

die Transformation dieser beiden Jordankastl

$$
\begin{aligned}
S\begin{pmatrix} J_j(\gamma) & 0 \\ 0 & J_j(\overline{\gamma}) \end{pmatrix} S^{-1} &= \frac{1}{2}\begin{pmatrix} E_j & E_j \\ -iE_j & iE_j \end{pmatrix}\begin{pmatrix} J_j(\gamma) & 0 \\ 0 & J_j(\overline{\gamma}) \end{pmatrix}\begin{pmatrix} E_j & iE_j \\ E_j & -iE_j \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} E_j & E_j \\ -iE_j & iE_j \end{pmatrix}\begin{pmatrix} J_j(\gamma) & iJ_j(\gamma) \\ J_j(\overline{\gamma}) & -iJ_j(\overline{\gamma}) \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} J_j(\gamma) + J_j(\overline{\gamma}) & iJ_j(\gamma) - iJ_j(\overline{\gamma}) \\ -iJ_j(\gamma) + iJ_j(\overline{\gamma}) & J_j(\gamma) + J_j(\overline{\gamma}) \end{pmatrix} \\
&= \begin{pmatrix} J_j(\mathrm{Re}\,(\gamma)) & -\mathrm{Im}\,(\gamma)E_j \\ \mathrm{Im}\,(\gamma)E_j & J_j(\mathrm{Re}\,(\gamma)) \end{pmatrix} =: B \in M_{2j}(\mathbb{R}) \qquad (11.3.27)
\end{aligned}
$$

und das gibt dann die sogenannte **reelle Jordan Normalform**, die aus zwei Jordankastl $J_j(\gamma), J_j(\overline{\gamma}) \in M_j(\mathbb{C})$ derselben Gestalt zu komplex konjugierten Eigenwerten einen reellen Block in $M_{2j}(\mathbb{R})$ macht. Die reelle Jordan Normalform ist zwar keine rechte obere Dreiecksmatrix wie die Jordan Normalform – das ist bei einem charakteristischem Polynom wie in (11.3.25) nach Satz 11.2.1 nicht möglich, aber eine Blockdiagonalmatrix.

## 11.4 e hoch Matrix

**Satz 11.4.1.** *Es sei* $\begin{array}{rcl} \|\cdot\|:\mathbb{K}^d & \to & [0,\infty[ \\ \underline{x} & \mapsto & \|\underline{x}\| \end{array}$ *eine Norm auf* $\mathbb{K}^d$ *und* $A \in M_d(\mathbb{K})$ *eine* $d\times d-$*Matrix mit Einträgen in* $\mathbb{K}$. *Dann definiert*

$$\begin{array}{rcl} |||\cdot|||: M_d(\mathbb{K}) & \to & [0,\infty[ \\ A & \mapsto & |||A||| := \sup\{\|A(\underline{x})\| : \underline{x}\in\mathbb{K}^d, \|\underline{x}\|\le 1\} \end{array} \tag{11.4.1}$$

*eine Norm auf* $M_d(\mathbb{K})$. *Diese heißt die von* $\|\cdot\|$ **induzierte Matrixnorm**. *Für* $\underline{x}\in\mathbb{K}^d$ *und* $A,B\in M_d(\mathbb{K})$ *gilt:*

$$\|A(\underline{x})\| \le |||A|||\,\|\underline{x}\| \tag{11.4.2}$$

$$|||AB||| \le |||A|||\,|||B||| \tag{11.4.3}$$

**Definition 11.4.2.** *Zu* $A\in M_d(\mathbb{K})$ *definiere* $A^0 := \mathrm{id}_{\mathbb{K}^d}$ *und*

$$e^A := \sum_{j=0}^{\infty} \frac{A^j}{j!} \tag{11.4.4}$$

**Bemerkung 11.4.3.** Laut (11.4.3) gilt $|||A^j||| \le |||A|||^j$ für eine induzierte Matrixnorm $|||\cdot|||$, $A\in M_d(\mathbb{K})$ und $j\in N$. Daher konvergiert

$$\sum_{j=0}^{\infty}\frac{|||A^j|||}{j!} \le \sum_{j=0}^{\infty}\frac{|||A|||^j}{j!} = e^{|||A|||},$$

also ist die Reihe $\displaystyle\sum_{j=0}^{\infty}\frac{A^j}{j!}$ absolut konvergent in der induzierten Matrixnorm $|||\cdot|||$. Wir werden im Kapitel über Topologie noch sehen, daß auf $\mathbb{K}^d$ alle Normen vollständig sind, ebenso die induzierten Matrixnormen, weshalb dann auch $e^A$ wohldefiniert ist.

**Bemerkung 11.4.4.** An dieser Stelle ist eine Warnung angebracht! Auch wenn es hier so scheint, als könne man $e^{tA}$ genauso behandeln, wie die komplexe Exponentialfunktion, ist dem nicht so: Beispielsweise gilt nur für $A,B\in M_d(\mathbb{K})$ mit $AB=BA$ die vom Komplexen geläufige Rechenregel $e^{A+B}=e^Ae^B$. Das wird aber falsch, wenn $AB\ne BA$ ist, wie etwa für $A=\begin{pmatrix}0&1\\0&0\end{pmatrix}$ und $B=\begin{pmatrix}0&0\\0&1\end{pmatrix}$ folgt. [3]

Wir halten jedoch noch mal ganz explizit fest:

[3] Hier ist $A^2=0$, $B^2=B$ und $(A+B)^2=A+B$, also $B^k=B$ und $(A+B)^k=A+B$ für jedes $k\in\mathbb{N}$, daher

$$\begin{aligned} e^{tA} &= \sum_{k=0}^{\infty}\frac{(tA)^k}{k!} = E_2+tA = \begin{pmatrix}1&t\\0&1\end{pmatrix} \\ e^{tB} &= \sum_{k=0}^{\infty}\frac{(tB)^k}{k!} = E_2 + B\sum_{k=1}^{\infty}\frac{t^k}{k!} = \begin{pmatrix}1&0\\0&e^t\end{pmatrix} \\ e^{t(A+B)} &= \sum_{k=0}^{\infty}\frac{(t(A+B))^k}{k!} = E_2+(A+B)\sum_{k=1}^{\infty}\frac{t^k}{k!} = \begin{pmatrix}1&e^t-1\\0&e^t\end{pmatrix} \ne \begin{pmatrix}1&t\\0&1\end{pmatrix}\cdot\begin{pmatrix}1&0\\0&e^t\end{pmatrix} \end{aligned}$$

**Satz 11.4.5.** *Es seien $A, B \in M_d(\mathbb{K})$ mit $AB = BA$, dann gilt $e^{A+B} = e^A e^B$. Insbesondere ist $e^0 = E_d \equiv \mathrm{id}_{\mathbb{K}^d}$ und $(e^{At})^{-1} = e^{-At}$ für alle $t \in \mathbb{R}$.*

*Beweis.* Da $AB = BA$ ist, und $e^A$ eine normkonvergente Potenzreihe ist, läßt sich der Beweis von $e^{z+w} = e^w e^z$ für die reelle/komplexe Exponentialfunktion mittels Cauchyprodukt sofort übertragen. □

**Satz 11.4.6.** *Es sei $A, T \in M_d(\mathbb{K})$ mit $detT \neq 0$ und*

$$J = T^{-1}AT = \begin{pmatrix} J_{d_1}(\alpha_1) & & & \\ & J_{d_2}(\alpha_2) & & \\ & & \ddots & \\ & & & J_{d_k}(\alpha_k) \end{pmatrix}$$

*sei von der Jordan Normalform wie in Satz 11.3.9 mit Jordankästchen*

$$J_{d_j}(\alpha_j) = \begin{pmatrix} \alpha_j & 1 & 0 & & \\ 0 & \alpha_j & 1 & \ddots & \\ & \ddots & \ddots & \ddots & 0 \\ & & \ddots & \ddots & 1 \\ & & & 0 & \alpha_j \end{pmatrix} \in M_{d_j}(\mathbb{K})$$

*Dann ist*

$$e^{tA} = Te^{tJ}T^{-1}, \tag{11.4.5}$$

*wobei $e^{tJ}$ wieder eine Blockmatrix der Gestalt*

$$e^{tJ} = \begin{pmatrix} e^{tJ_{d_1}(\alpha_1)} & & & \\ & e^{tJ_{d_2}(\alpha_2)} & & \\ & & \ddots & \\ & & & e^{tJ_{d_k}(\alpha_k)} \end{pmatrix} \tag{11.4.6}$$

*ist, und "e hoch t mal Jordankästchen" hat die Form*

$$e^{tJ_{d_j}(\alpha_j)} = e^{t\alpha_j} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \frac{t^3}{3!} & \cdots & \cdots & \frac{t^{d_j-1}}{(d_j-1)!} \\ 0 & 1 & t & \frac{t^2}{2!} & \cdots & \cdots & \frac{t^{d_j-2}}{(d_j-2)!} \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \\ \vdots & & & \ddots & \ddots & \ddots & \frac{t^2}{2!} \\ \vdots & & & & \ddots & \ddots & t \\ 0 & & & & & 0 & 1 \end{pmatrix} \in M_{d_j}(\mathbb{K}). \tag{11.4.7}$$

*Beweis.* Wegen $A = TJT^{-1}$ bekommen wir $A^m = (TJT^{-1})^m = TJ^mT^{-1}$. also

$$e^{tA} = \sum_{m=0}^{\infty} \frac{t^m}{m!} A^m = \sum_{m=0}^{\infty} \frac{t^m}{m!} (TAT^{-1})^m = T \sum_{m=0}^{\infty} \frac{t^m}{m!} J^m T^{-1} = Te^{tJ}T^{-1}.$$

$J$ ist eine Matrix in Jordan Normalform, bei der Bildung von Potenzen $J^m$ bleibt aufgrund der Multiplikationsregel für Blockmatrizen die Blockmatrixstruktur erhalten:

$$J^m = J = \begin{pmatrix} J_{d_1}(\alpha_1)^m & & & \\ & J_{d_2}(\alpha_2)^m & & \\ & & \ddots & \\ & & & J_{d_k}(\alpha_k)^m \end{pmatrix}, \tag{11.4.8}$$

mit $J_{d_1}(\alpha_1)^m \in M_{d_1}(\mathbb{K}), ..., J_{d_k}(\alpha_k)^m \in M_{d_k}(\mathbb{K})$, folglich ist

$$e^{tJ} = \begin{pmatrix} e^{tJ_{d_1}(\alpha_1)} & & & \\ & e^{tJ_{d_2}(\alpha_2)} & & \\ & & \ddots & \\ & & & e^{tJ_{d_k}(\alpha_k)} \end{pmatrix} \tag{11.4.9}$$

Zur Berechnung von $e^{tJ_{d_j}(\alpha_j)}$ beachte, daß für $J_{d_j}(0) = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \in M_{d_j}(\mathbb{K})$ gilt:

$$J_{d_j}(0)^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & & & \\ & 0 & 0 & 1 & 0 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots & 0 \\ & & & & \ddots & \ddots & 1 \\ & & & & & \ddots & 0 \\ & & & & & & 0 \end{pmatrix}, \qquad J_{d_j}(0)^3 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & & & \\ & 0 & 0 & 0 & 1 & 0 & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots & \ddots & 0 \\ & & & & \ddots & \ddots & \ddots & 1 \\ & & & & & \ddots & \ddots & 0 \\ & & & & & & \ddots & 0 \\ & & & & & & & 0 \end{pmatrix},$$

$$\cdots, J_{d_j}(0)^{d_j-1} = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ & 0 & \cdots & 0 \\ & & \ddots & \vdots \\ & & & 0 \end{pmatrix}, \qquad J_{d_j}(0)^{d_j} = 0 \in M_{d_j}(\mathbb{K})$$

$E_{d_j}$ und $J_{d_j}(0)$ kommutieren, deshalb gilt nach Satz 11.4.5:

$$e^{tJ_{d_j}(\alpha_j)} = e^{t\alpha_j E_{d_j} + tJ_{d_j}(0)} = e^{t\alpha_j} E_{d_j} e^{tJ_{d_j}(0)} = e^{t\alpha_j} \sum_{m=0}^{\infty} \frac{t^m}{m!} J_{d_j}(0)^m = e^{t\alpha_j} \sum_{m=0}^{d_j-1} \frac{t^m}{m!} J_{d_j}(0)^m.$$

□

**Korollar 11.4.7.** *Es sei $A \in M_d(\mathbb{K})$ diagonalisierbar und* $T^{-1}AT = \begin{pmatrix} \nu_1 & & & \\ & \nu_2 & & \\ & & \ddots & \\ & & & \nu_d \end{pmatrix}$.

*Dann ist*

$$e^{tA} = T \begin{pmatrix} e^{t\nu_1} & & & \\ & e^{t\nu_2} & & \\ & & \ddots & \\ & & & e^{t\nu_d} \end{pmatrix} T^{-1}. \tag{11.4.10}$$

*Für jeden Eigenvektor $v_j$ von $A$ zum Eigenwert $\nu_j$ gilt $e^{tA}v_j = e^{t\nu_j}v_j$.*

**Beispiel 11.4.8.** In Beispiel 11.3.13 hatten wir für

$$A = \begin{pmatrix} 0 & -1 & -2 \\ -1 & 0 & 2 \\ 1 & -1 & -3 \end{pmatrix}$$

die Jordanform $J = T^{-1}AT = \begin{pmatrix} -1 & 1 & \\ & -1 & \\ & & -1 \end{pmatrix}$ mit den Transformationsmatrizen $T = \begin{pmatrix} -2 & 0 & 1 \\ 2 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}$ und $T^{-1} = \frac{1}{4}\begin{pmatrix} -1 & 1 & 0 \\ -2 & 2 & 4 \\ 2 & 2 & 0 \end{pmatrix}$ bestimmt. Wir berechnen jetzt $e^{tA}$ nach Satz 11.4.6:

$$e^{tJ} = e^{-t}\begin{pmatrix} 1 & t & \\ & 1 & \\ & & 1 \end{pmatrix}$$

und damit folgt wegen $A = TJT^{-1}$

$$\begin{aligned} e^{tA} &= Te^{tJ}T^{-1} = \begin{pmatrix} -2 & 0 & 1 \\ 2 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix} e^{-t}\begin{pmatrix} 1 & t & \\ & 1 & \\ & & 1 \end{pmatrix}\frac{1}{4}\begin{pmatrix} -1 & 1 & 0 \\ -2 & 2 & 4 \\ 2 & 2 & 0 \end{pmatrix} \\ &= e^{-t}\begin{pmatrix} 1+t & -t & -2t \\ -t & 1+t & 2t \\ t & -t & -2t+1 \end{pmatrix}. \end{aligned}$$

# Kapitel 12

# Selbstadjungierte und unitäre Matrizen

## 12.1 Winkel, orthogonale Vektoren und Orthonormalbasen

**Definition 12.1.1.** *Sei $(V, \langle \cdot, \cdot \rangle)$ ein $\mathbb{K}$-Vektorraum mit Skalarprodukt.*

1. *$v, w \in V$ heissen orthogonal (in Zeichen $v \perp w$), wenn $\langle v, w \rangle = 0$.*

2. *$v_1, ..., v_n \in V$ heissen orthonormal, wenn $\|v_1\| = ... = \|v_n\| = 1$ und $v_i \perp v_j$ für $1 \leq i < j \leq n$.*

3. *Ist $V$ ein euklidischer Vektorraum und sind $v, w \in V \backslash \{\mathbf{0}\}$, so definiert man durch*

$$\cos(\measuredangle(v, w)) = \frac{\langle v, w \rangle}{\|v\| \|w\|} \in [-1, 1] \tag{12.1.1}$$

   *den Winkel $\measuredangle(v, w) \in [0, \pi]$ zwischen $v$ und $w$.*

**Lemma 12.1.2.** *Es sei $V$ ein $\mathbb{K}$-Vektorraum mit Skalarprodukt und $v_1, ..., v_n \in V \backslash \{\mathbf{0}\}$ seien paarweise orthogonale Vektoren. Dann sind $v_1, ..., v_n$ linear unabhängig.*

*Beweis.* Sind $\lambda_1, ..., \lambda_n \in \mathbb{K}$ mit $\lambda_1 v_1 + ... + \lambda_n v_n = \mathbf{0}$, so folgt für jedes $j \in \{1, ..., n\}$ durch Bilden des Skalarprodukts:

$$0 = \langle v_j, \lambda_1 v_1 + ... + \lambda_n v_n \rangle = \lambda_1 \langle v_j, v_1 \rangle + ... + \lambda_n \langle v_j, v_n \rangle = \lambda_j \langle v_j, v_j \rangle$$

und da $v_j \neq \mathbf{0}$ ist, gilt $\langle v_j, v_j \rangle \neq 0$, daher muß $\lambda_j = 0$ sein. $\square$

**Definition 12.1.3.** *Es sei $V$ ein $n$-dimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt, dann heißt eine Menge $\{v_1, ..., v_n\}$ bestehend aus $n$ orthonormalen Vektoren eine* **Orthonormalbasis** *von $V$.*

**Beispiel 12.1.4.**

1. Betrachte $\mathbb{R}^3$ mit dem Standardskalarprodukt. Im Beispiel 11.3.13 haben wir die Basis $\mathcal{B}_1 = \left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \right\}$ von $\mathbb{R}^3$ als Basis aus Eigenvektoren der Matrix $A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ -2 & -3 & 0 \end{pmatrix}$ ermittelt. Aber dies ist keine Orthonormalbasis, denn zum Beispiel ist $\left\langle \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \right\rangle_{\mathbb{R}^3} = -4$ oder $\left\| \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\| = \sqrt{2}$.

2. $\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ ist eine Orthonormalbasis von $(\mathbb{R}^3, \langle \cdot, \cdot \rangle_{\mathbb{R}^3})$.

**Satz 12.1.5** (Schmidt'sches Orthonormalisierungsverfahren)**.** *Es sei $V$ ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $w_1, ..., w_m$ seien orthonormal und $v \notin \operatorname{lin}\{w_1, ..., w_m\}$. Dann bilden für*

$$w := v - \langle w_1, v \rangle w_1 - \langle w_2, v \rangle w_2 - ... - \langle w_m, v \rangle w_m \tag{12.1.2}$$

*und $w_{m+1} := \frac{w}{\|w\|}$ die Vektoren $w_1, ..., w_m, w_{m+1}$ eine Familie von orthonormalen Vektoren mit $\operatorname{lin}\{w_1, ..., w_m, v\} = \operatorname{lin}\{w_1, ..., w_m, w_{m+1}\}$.*

*Beweis.* Aus (12.1.2) folgt

$$\langle w_k, w \rangle = \langle w_k, v \rangle - \left(\sum_{j=1}^{m} \langle w_j, v \rangle \langle w_k, w_j \rangle\right) = \langle w_k, v \rangle - \langle w_k, v \rangle = 0$$

für alle $k \in \{1, ..., m\}$. Da nach Voraussetzung $w_1, ..., w_m, v$ linear unabhängig sind, ist $w \neq \mathbf{0}$, daher können wir $w_{m+1} := \dfrac{w}{\|w\|}$ definieren und somit sind $w_1, ..., w_m, w_{m+1}$ orthonormal. Nach Definition von $w$ bzw. $w_{m+1}$ folgt $\operatorname{lin}\{w_1, ..., w_m, v\} = \operatorname{lin}\{w_1, ..., w_m, w_{m+1}\}$. □

**Korollar 12.1.6.** *Jeder $n$-dimensionale $\mathbb{K}$-Vektorraum $V \neq \{\mathbf{0}\}$ besitzt eine Orthonormalbasis.*

**Beispiel 12.1.7.** Es sei $V := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4 : x_1 + 2x_2 - x_3 + x_4 = 0 \right\}$, dann ist offenbar $\dim_{\mathbb{R}}(V) = 3$ und $\mathcal{A} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}$ ist ein linear unabhängiges System in $V$, also eine Basis von $V$. Mit dem Schmidt'schen Orthonormalisierungsverfahren bekommen wir daraus folgende Orthonormalbasis von $V$:

- Normiere einen der Basisvektoren zB. $w_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$.

- Wende beginnend mit dem System $\{w_1\}$ orthonormaler Vektoren sukzessive das Orthonormalisierungsverfahren an: Da $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}$ linear unabhängig sind, betrachte im ersten Schritt

$$w := \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}} \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}$$

und normiere diesen Vektor zu $w_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}$, dann bilden $\{w_1, w_2\}$ eine Orthonormalbasis von $\operatorname{lin} \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}$.

- Für $W = \operatorname{lin}\{w_1, w_2\}$ betrachte

$$\begin{aligned} w \quad := \quad & \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}} \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\rangle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \\ & - \frac{1}{\sqrt{6}} \left\langle \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\rangle \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 \\ -3 \\ -2 \\ 2 \end{pmatrix} \end{aligned}$$

und normiere diesen zu $w_3 = \frac{1}{\sqrt{21}} \begin{pmatrix} 2 \\ -3 \\ -2 \\ 2 \end{pmatrix}$, dann bildet $\{w_1, w_2, w_3\}$ eine Orthonormalbasis von $V$.

## 12.2 Selbstadjungierte Endomorphismen und Matrizen

**Definition 12.2.1.** *Es sei*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{pmatrix} \in M(m \times n, \mathbb{K}),$$

*dann heißt*

$$A^* := \begin{pmatrix} \overline{a_{11}} & \overline{a_{21}} & \cdots & \overline{a_{m1}} \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & \overline{a_{m2}} \\ \vdots & & & \vdots \\ \overline{a_{1n}} & \cdots & \cdots & \overline{a_{mn}} \end{pmatrix} \in M(n \times m, \mathbb{K}),$$

*die* **adjungierte Matrix** *zu $A$. Eine Matrix $A \in M_n(\mathbb{K})$ mit $A = A^*$ heißt* **selbstadjungiert**.

**Lemma 12.2.2.** *Ist $A \in M(m \times n, \mathbb{K})$ und $\langle \cdot, \cdot \rangle_{\mathbb{K}^n}$ das Standardskalarprodukt auf $\mathbb{K}^n$, so gilt:*

$$\langle \underline{x}, A\underline{y} \rangle_{\mathbb{K}^m} = \langle A^* \underline{x}, \underline{y} \rangle_{\mathbb{K}^n}$$

*für alle $\underline{x} \in \mathbb{K}^m, \underline{y} \in \mathbb{K}^n$.*

*Beweis.* Nach den Regeln für die Matrixmultiplikation und das Standardskalarprodukt gilt:

$$\begin{aligned} \langle \underline{x}, A\underline{y} \rangle_{\mathbb{K}^m} &= \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \begin{pmatrix} \sum_{l=1}^{n} a_{1l} y_l \\ \vdots \\ \sum_{l=1}^{n} a_{ml} y_l \end{pmatrix} \right\rangle = \sum_{k=1}^{m} \sum_{l=1}^{n} \overline{x_k} a_{kl} y_l = \\ \langle A^* \underline{x}, \underline{y} \rangle_{\mathbb{K}^n} &= \left\langle \begin{pmatrix} \sum_{k=1}^{m} \overline{a_{k1}} x_k \\ \vdots \\ \sum_{k=1}^{m} \overline{a_{kn}} x_k \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{l=1}^{n} \overline{\left( \sum_{k=1}^{m} \overline{a_{kl}} x_k \right)} y_l \end{aligned}$$

$\square$

**Definition 12.2.3.** *Es sei $V$ ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Eine $\mathbb{K}$-lineare Abbildung $F : V \to V$ heißt* **selbstadjungiert***, wenn für alle $v, w \in V$ gilt:*

$$\langle F(v), w \rangle = \langle v, F(w) \rangle. \tag{12.2.1}$$

**Lemma 12.2.4.** *Alle Eigenwerte einer selbstadjungierten Abbildung sind reell.*

*Beweis.* Ist $\lambda \in \mathbb{K}$ Eigenwert der selbstadjungierten Abbildung $F : V \longrightarrow V$, so gilt für jeden Eigenvektor $v \neq 0$ von $F$ zum Eigenwert $\lambda$:

$$\overline{\lambda}\langle v, v\rangle = \langle \lambda v, v\rangle = \langle F(v), v\rangle = \langle v, F(v)\rangle = \langle v, \lambda v\rangle = \lambda\langle v, v\rangle,$$

also ist $\lambda = \overline{\lambda}$ oder $\lambda \in \mathbb{R}$. □

**Lemma 12.2.5.** *Es sei $(V, \langle\cdot,\cdot\rangle)$ ein $n$-dimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt und $F : V \longrightarrow V$ sei $\mathbb{K}$-linear, $\mathcal{B}$ eine Orthonormalbasis von $V$ und $M_{\mathcal{B}}^{\mathcal{B}}(F)$ sei die darstellende Matrix von $F$ bezüglich $\mathcal{B}$. Dann sind äquivalent:*

*a) $F$ ist selbstadjungiert*

*b) $M_{\mathcal{B}}^{\mathcal{B}}(F)$ ist selbstadjungiert.*

*Beweis.* Zu $v, w \in V$ seien $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\underline{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{K}^n$ die Koordinatenvektoren bezüglich $\mathcal{B} = (v_1, ..., v_n)$, also $v = x_1v_1 + ... + x_nv_n$ und $w = y_1v_1 + ... + y_nv_n$. Aus den Eigenschaften des Skalarprodukts folgt

$$\langle v, w\rangle = \sum_{i,j=1}^{n} \overline{x_i} y_j \langle v_i, v_j\rangle,$$

also ist $\mathcal{B}$ genau dann eine Orthonormalbasis von $(V, \langle\cdot,\cdot\rangle)$, wenn $\langle v, w\rangle = \langle \underline{x}, \underline{y}\rangle_{\mathbb{K}^n}$ für alle $v, w \in V$ und den entsprechenden Koordinatenvektoren $\underline{x}, \underline{y} \in \mathbb{K}^n$ gilt. Die Selbstadjungiertheit von $F$ ist also äquivalent zur Forderung

$$\langle M_{\mathcal{B}}^{\mathcal{B}}(F)\underline{x}, \underline{y}\rangle_{\mathbb{K}^n} = \langle \underline{x}, M_{\mathcal{B}}^{\mathcal{B}}(F)\underline{y}\rangle_{\mathbb{K}^n} \qquad \text{für alle } \underline{x}, \underline{y} \in \mathbb{K}^n \tag{12.2.2}$$

an die darstellende Matrix. Nach Lemma 12.2.2 ist $\langle M_{\mathcal{B}}^{\mathcal{B}}(F)\underline{x}, \underline{y}\rangle_{\mathbb{K}^n} = \langle \underline{x}, M_{\mathcal{B}}^{\mathcal{B}}(F)^*\underline{y}\rangle_{\mathbb{K}^n}$ für alle $\underline{x}, \underline{y} \in \mathbb{K}^n$ und daher ist (12.2.2) nach Lemma 5.4.15 äquivalent zu $M_{\mathcal{B}}^{\mathcal{B}}(F)\underline{y} = M_{\mathcal{B}}^{\mathcal{B}}(F)^*\underline{y}$ für alle $\underline{y} \in \mathbb{K}^n$, was wieder $M_{\mathcal{B}}^{\mathcal{B}}(F) = M_{\mathcal{B}}^{\mathcal{B}}(F)^*$ zeigt. □

**Lemma 12.2.6.** *Bei jeder selbstadjungierten Abbildung $F : V \to V$ sind Eigenvektoren zu verschiedenen Eigenwerten orthogonal.*

*Beweis.* Für selbstadjungiertes $F$ sind alle Eigenwerte nach Lemma 12.2.4 reell. Sind also $\lambda_1 \neq \lambda_2$ Eigenwerte von $F$ und $v_1$ Eigenvektor von $F$ zum Eigenwert $\lambda_1$ und $v_2$ Eigenvektor von $F$ zum Eigenwert $\lambda_2$, so ist

$$0 = \langle F(v_1), v_2\rangle - \langle v_1, F(v_2)\rangle = \langle \lambda_1 v_1, v_2\rangle - \langle v_1, \lambda_2 v_2\rangle = (\lambda_1 - \lambda_2)\langle v_1, v_2\rangle,$$

und wegen $\lambda_1 \neq \lambda_2$ ist $\langle v_1, v_2\rangle = 0$ . □

**Satz 12.2.7.** *Ist $V$ ein $n$-dimensionaler $\mathbb{K}-$ Vektorraum mit Skalarprodukt und $F : V \to V$ selbstadjungiert, so besitzt $V$ eine Orthonormalbasis aus Eigenvektoren von $F$.*

*Beweis.* Das charakteristische Polynom $p_F = (\lambda_1 - X)\cdots(\lambda_n - X)$ zerfällt über $\mathbb{C}$ nach dem Fundamentalsatz der Algebra in Linearfaktoren. Dabei gilt nach Lemma 12.2.4 sogar $\lambda_1, ..., \lambda_n \in \mathbb{R}$. Wir zeigen jetzt die Behauptung durch Induktion nach $n = \dim V$:

- Für $n = 1$ bildet jeder Vektor $v \in V$ mit $\|v\| = 1$ eine Ortonormalbasis von $V = \mathrm{Eig}(F, \lambda_1)$.
- Induktionsschritt: $n - 1 \to n$: Es sei $F(v_1) = \lambda_1 v_1$ mit $\|v_1\| = 1$; definiere nun

  $$W := \{w \in V : \langle w, v_1\rangle = 0\},$$

  dann ist $\langle F(w), v_1\rangle = \langle w, F(v_1)\rangle = \langle w, \lambda_1 v_1\rangle = \lambda_1\langle w, v_1\rangle = 0$ für $w \in W$, also $F(W) \subseteq W$. Da $F$ selbstadjungiert ist, ist auch die Einschränkung $G := F|_W : W \longrightarrow W$ selbstadjungiert. Wegen $\dim(W) = n - 1$ gibt [1] es nach Induktionsannahme eine Orthonormalbasis $\mathcal{B}_W = \{v_2, ..., v_n\}$ von $W$ bestehend aus Eigenvektoren von $G$. Nach Definition von $W$ gilt $\langle v_1, v_k\rangle = 0$ für alle $k \in \{2, ..., n\}$ und damit ist $\mathcal{B} = \{v_1, ..., v_n\}$ eine Orthonormalbasis von $V$ aus Eigenvektoren von $F$.

□

**Rezept 12.2.8** (zur Bestimmung einer Orthonormalbasis aus Eigenvektoren eines selbstadjungierten Endomorphismus)**.** Ist $V$ ein $n$-dimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt und $F : V \to V$ ein selbstadjungierter Endomorphismus, so

- bestimme das charakteristische Polynom

  $$p_F = (\alpha_1 - X)^{a_1}\cdots(\alpha_m - X)^{a_m}$$

  von $F$ – hier ist $\alpha_1, ..., \alpha_m \in \mathbb{R}$, $a_1, ..., a_m \in \mathbb{N}$ mit $a_1 + ...a_m = n$.
- berechne für alle Eigenwerte $\alpha_1, ..., \alpha_m$ die Eigenräume $\mathrm{Eig}(F, \alpha_1), ..., \mathrm{Eig}(F, \alpha_m)$
- damit bekommt man für jeden Eigenraum $\mathrm{Eig}(F, \alpha_k)$ eine Basis $\mathcal{A}_k$, $k = 1, ..., m$.
- wende das Schmidtsche Orthonormalisierungsverfahren auf die Basis $\mathcal{A}_k$ von $\mathrm{Eig}(F, \alpha_k)$ an und erhalte eine Orthonormalbasis $\mathcal{B}_k$ von $\mathrm{Eig}(F, \alpha_k)$
- die Vereinigung $\mathcal{B} = \bigcup_{k=1}^{m} \mathcal{B}_k$ ist dann eine Orthonormalbasis von $V$ bestehend aus Eigenvektoren von $F$.

---

[1] $W = \mathrm{Kern}(h)$ mit $\begin{array}{rcl} h : V & \to & \mathbb{K} \\ w & \mapsto & \langle v_1, w\rangle \end{array}$ und wegen $h(v_1) = \|v_1\|^2 = 1$ ist der Untervektorraum $\{0\} \subsetneq h(V) \subseteq \mathbb{K}$, also $h(V) = \mathbb{K}$ und daher folgt nach der Dimensionsformel:

$$\dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(\mathrm{Kern}(h)) = \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(h(V)) = n - 1.$$

**Beispiel 12.2.9.** Wir betrachten $A = \frac{1}{15}\begin{pmatrix} 10 & 5 & 10 \\ 5 & -14 & 2 \\ 10 & 2 & -11 \end{pmatrix}$ und untersuchen, welche Orthonormalbasis es für $\mathbb{R}^3$ aus Eigenvektoren von $A$ gibt.

$$\begin{aligned}
p_A &= \det(A - XE_3) = \det\begin{pmatrix} \frac{2}{3} - X & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{14}{15} - X & \frac{2}{15} \\ \frac{2}{3} & \frac{2}{15} & -\frac{11}{15} - X \end{pmatrix} = \\
&= (\frac{2}{3} - X)(\frac{14}{15} + X)(\frac{11}{15} + X) + \frac{8}{135} + \frac{4}{9}(\frac{14}{15} + X) - \frac{4}{225}(\frac{2}{3} - X) + \frac{1}{9}(\frac{11}{15} + X) \\
&= 1 + X - X^2 - X^3 = -(X-1)(X+1)^2
\end{aligned}$$

ist die Zerlegung für das charakteristische Polynom von $A$, also sind $\lambda_1 = 1$ und $\lambda_2 = -1$ die Eigenwerte von $A$. Zur Bestimmung von $\mathrm{Eig}(A, 1)$ forme $A - E_3$ durch elementare Zeilenumformungen auf Zeilenstufenform um:

$$\begin{aligned}
A - E_3 &= \frac{1}{15}\begin{pmatrix} -5 & 5 & 10 \\ 5 & -29 & 2 \\ 10 & 2 & -26 \end{pmatrix} \xrightarrow{(II)+(I),(III)+2(I)} \frac{1}{15}\begin{pmatrix} -5 & 5 & 10 \\ 0 & -24 & 12 \\ 0 & 12 & -6 \end{pmatrix} \longrightarrow \\
&\xrightarrow{(III)+\frac{1}{2}(II)} \frac{1}{15}\begin{pmatrix} -5 & 5 & 10 \\ 0 & -24 & 12 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{3(I),\frac{5}{4}(II)} \begin{pmatrix} -1 & 1 & 2 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix}
\end{aligned}$$

Also ist $\mathrm{Eig}(A, 1)$ eindimensional und $v_1 := \frac{1}{\sqrt{30}}\begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix}$ einer der normierten Eigenvektoren in $\mathrm{Eig}(A, 1)$. Bestimmen von $\mathrm{Eig}(A, -1)$:

$$A + E_3 = \frac{1}{15}\begin{pmatrix} 25 & 5 & 10 \\ 5 & 1 & 2 \\ 10 & 2 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 5 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

somit ist $\mathrm{Eig}(A, -1)$ zweidimensional. Die beiden Eigenvektoren $\begin{pmatrix} -1 \\ 5 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix}$ sind linear unabhängig. $w_1 := \frac{1}{\sqrt{26}}\begin{pmatrix} -1 \\ 5 \\ 0 \end{pmatrix}$ ist der Einheitsvektor zum ersten Vektor und das Schmidtsche Orthonormalisierungsverfahren liefert

$$\begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix} - \frac{1}{26}\left\langle \begin{pmatrix} -1 \\ 5 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix} \right\rangle \begin{pmatrix} -1 \\ 5 \\ 0 \end{pmatrix} = \frac{5}{13}\begin{pmatrix} -5 \\ -1 \\ 13 \end{pmatrix}$$

so daß $w_1$ zusammen mit $w_2 := \frac{1}{\sqrt{195}}\begin{pmatrix} -5 \\ -1 \\ 13 \end{pmatrix}$ eine Orthonormalbasis von $\mathrm{Eig}(A, -1)$ bilden. $\{v_1, w_1, w_2\}$ ist jetzt eine Orthonormalbasis von $\mathbb{R}^3$ bestehend aus Eigenvektoren von $A$.

## 12.3 Orthogonale und unitäre Endomorphismen und Matrizen

**Definition 12.3.1.** *Eine $\mathbb{C}$-lineare Abbildung $F : V \longrightarrow V$ eines unitären Vektorraums $V$ heißt* **unitär**, *wenn für alle $v, w \in V$*

$$\langle F(v), F(w)\rangle = \langle v, w\rangle \tag{12.3.1}$$

*gilt. Eine Matrix $A \in M_n(\mathbb{C})$ heißt* **unitär**, *wenn $A$ invertierbar ist und $A^{-1} = A^*$ gilt. Eine Abbildung $F : V \to V$ eines euklidischen Vektorraums mit $\langle F(v), F(w)\rangle = \langle v, w\rangle$ für alle $v, w \in V$ heißt* **orthogonal**. *Eine invertierbare Matrix $A \in M_n(\mathbb{R})$ mit $A^{-1} = A^T$ heißt* **orthogonal**.

**Bemerkung 12.3.2.** Die Bedingung $\langle F(v), F(w)\rangle = \langle v, w\rangle$ bei orthogonalen Abbildungen zeigt uns in den leicht vorstellbaren Fällen $\mathbb{R}^2$ und $\mathbb{R}^3$, daß $F$ die Längen der Vektoren und die Winkel zwischen zwei Vektoren nicht verändert.

**Lemma 12.3.3.** *Ist $\lambda$ ein Eigenwert einer unitären/orthogonalen Abbildung $F$, dann ist $|\lambda| = 1$.*

*Beweis.* Es sei $v \neq 0$ ein Eigenvektor von $F$ zum Eigenwert $\lambda$, also $F(v) = \lambda v$, dann folgt:

$$\overline{\lambda}\lambda\langle v, v\rangle = \langle \lambda v, \lambda v\rangle = \langle F(v), F(v)\rangle = \langle v, v\rangle,$$

also ist $\overline{\lambda}\lambda = |\lambda|^2 = 1$. □

**Lemma 12.3.4.** *Es sei $V$ ein $n$-dimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt und $F : V \longrightarrow V$ sei $\mathbb{K}$-linear, $\mathcal{B}$ eine Orthonormalbasis von $V$ und $M^{\mathcal{B}}_{\mathcal{B}}(F)$ die darstellende Matrix von $F$ bezüglich $\mathcal{B}$. Dann sind äquivalent*

1. *$F$ ist unitär bzw. orthogonal*
2. *$M^{\mathcal{B}}_{\mathcal{B}}(F)$ ist unitär bzw. orthogonal.*

*Beweis.* Wie beim Beweis von Lemma 12.2.5 folgt für eine Orthonormalbasis $\mathcal{B}$ die Äquivalenz der Bedingungen $\langle F(v), F(w)\rangle = \langle v, w\rangle$ und $\langle M^{\mathcal{B}}_{\mathcal{B}}(F)\underline{x}, M^{\mathcal{B}}_{\mathcal{B}}(F)\underline{y}\rangle_{\mathbb{K}^n} = \langle \underline{x}, \underline{y}\rangle_{\mathbb{K}^n}$, wenn $\underline{x}, \underline{y} \in \mathbb{K}^n$ die Koordinatenvektoren von $v$ bzw. $w$ bezüglich der Basis $\mathcal{B}$ bezeichnen. Nach Lemma 12.2.2 ist dies zu $\langle \underline{x}, M^{\mathcal{B}}_{\mathcal{B}}(F)^* M^{\mathcal{B}}_{\mathcal{B}}(F)\underline{y}\rangle_{\mathbb{K}^n} = \langle \underline{x}, \underline{y}\rangle_{\mathbb{K}^n}$ äquivalent, was gemäß Lemma 5.4.15 zu $M^{\mathcal{B}}_{\mathcal{B}}(F)^* M^{\mathcal{B}}_{\mathcal{B}}(F) = E_n$ äquivalent ist. Damit hat $M^{\mathcal{B}}_{\mathcal{B}}(F)$ das Linksinverse $M^{\mathcal{B}}_{\mathcal{B}}(F)^*$ in $GL(n, \mathbb{K})$ und da $GL(n, \mathbb{K})$ eine Gruppe ist, folgt daraus $M^{\mathcal{B}}_{\mathcal{B}}(F)^* = M^{\mathcal{B}}_{\mathcal{B}}(F)^{-1}$, dh. $M^{\mathcal{B}}_{\mathcal{B}}(F)$ unitär bzw. orthogonal. □

Die Bedingung $A = A^*$ für selbstadjungierte Matrizen läßt sich leicht nachprüfen. Schwieriger wird dies mit den Bedingungen $A^{-1} = A^T$ für orthogonale oder $A^{-1} = A^*$ für unitäre Matrizen. Wir müssen aber nicht die inverse Matrix $A^{-1}$ berechnen um zu entscheiden, ob eine gegebene Matrix $A$ orthogonal oder unitär ist.

**Lemma 12.3.5.** *Für $A \in M_n(\mathbb{K})$ sind äquivalent*

*a) $A$ ist orthogonal / unitär.*

*b)* $A^*A = E_n$

*c) Die Spaltenvektoren von A sind eine Orthonormalbasis von* $(\mathbb{K}^n, \langle\cdot,\cdot\rangle_{\mathbb{K}^n})$.

*d) Die Zeilenvektoren von A sind eine Orthonormalbasis von* $(\mathbb{K}^n, \langle\cdot,\cdot\rangle_{\mathbb{K}^n})$.

*Beweis.* Ist $A$ unitär, also $A^{-1} = A^*$, so gilt $A^*A = E_n$. Ist umgekehrt $A^*A = E_n$, so hat $A$ in $GL(n,\mathbb{K})$ das linksinverse Element $A^*$ und da $GL(n,\mathbb{K})$ eine Gruppe ist, folgt $A^{-1} = A^*$, was die Äquivalenz von a) und b) zeigt.

Ist $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = (\underline{a}_1, ..., \underline{a}_n)$, dann ist $A^* = \begin{pmatrix} \overline{a_{11}} & \cdots & \overline{a_{n1}} \\ \vdots & & \vdots \\ \overline{a_{1n}} & \cdots & \overline{a_{nn}} \end{pmatrix}$ und

$$A^*A = \begin{pmatrix} \langle \underline{a}_1, \underline{a}_1\rangle_{\mathbb{K}^n} & \cdots & \langle \underline{a}_1, \underline{a}_n\rangle_{\mathbb{K}^n} \\ \vdots & & \vdots \\ \langle \underline{a}_n, \underline{a}_1\rangle_{\mathbb{K}^n} & \cdots & \langle \underline{a}_n, \underline{a}_n\rangle_{\mathbb{K}^n} \end{pmatrix},$$

was die Äquivalenz von b) und c) zeigt. Ist $A$ invertierbar, so ist es auch $A^T$ und es gilt $(A^T)^{-1} = (A^{-1})^T$, daher ist $A$ genau dann orthogonal/unitär, wenn $A^T$ orthogonal/unitär ist. Da die Zeilenvektoren von $A$ gerade die Spaltenvektoren von $A^T$ sind, folgt daraus die Äquivalenz von a) und d). □

**Lemma 12.3.6.** *Es sei V ein* $\mathbb{K}$-*Vektorraum mit Skalarprodukt und* $F : V \to V$ *sei orthogonal/unitär, dann sind Eigenvektoren zu verschiedenen Eigenwerten orthogonal.*

*Beweis.* Es seien $\lambda_1 \neq \lambda_2$ Eigenwerte von $F$ und $v_1, v_2 \in V\backslash\{\mathbf{0}\}$ mit $F(v_1) = \lambda_1 v_1$ und $F(v_2) = \lambda_2 v_2$. Dann ist

$$\langle v_1, v_2\rangle = \langle F(v_1), F(v_2)\rangle = \langle \lambda_1 v_1, \lambda_2 v_2\rangle = \overline{\lambda_1}\lambda_2\langle v_1, v_2\rangle$$

Angenommen $\langle v_1, v_2\rangle \neq 0$, dann ist $\overline{\lambda_1}\lambda_2 = 1$ und aus $|\lambda_1|^2 = 1$ folgt daraus $\lambda_1 = \lambda_1(\overline{\lambda_1}\lambda_2) = (\lambda_1\overline{\lambda_1})\lambda_2 = \lambda_2$ im Widerspruch zur Voraussetzung. □

**Satz 12.3.7.** *Ist V ein unitärer Vektorraum mit* $\dim_{\mathbb{C}} V = n < \infty$ *und* $F : V \longrightarrow V$ *unitär, so besitzt V eine Orthonormalbasis aus Eigenvektoren von F.*

*Beweis.* Nach dem Fundamentalsatz der Algebra schreibt sich das charakteristische Polynom

$$p_F = (\lambda_1 - X)\cdots(\lambda_n - X)$$

von $F$ mit – nicht notwendigerweise verschiedenen – $\lambda_1, ..., \lambda_n \in \mathbb{C}$ und in Lemma 12.3.3 haben wir $|\lambda_1| = ... = |\lambda_n| = 1$ bewiesen. Wir zeigen die Behauptung durch Induktion nach $n$:

- Induktionsanfang $n = 1$: Dann ist $V = \mathrm{Eig}(F, \lambda_1)$ und jeder Einheitsvektor $v_1 \in V$ bildet eine Orthonormalbasis von $V$ aus Eigenvektoren von $F$.

- Induktionsschritt $n-1 \to n$. Ist nun $v_1$ Eigenvektor von $F$ zum Eigenwert $\lambda_1$ mit $\|v_1\| = 1$, so definieren wir

  $$W := \{w \in V : \langle w, v_1\rangle = 0\},$$

  dann gilt

  $$\lambda_1\langle F(w), v_1\rangle = \langle F(w), \lambda_1 v_1\rangle = \langle F(w), F(v_1)\rangle = \langle w, v_1\rangle = 0$$

  für jedes $w \in W$ und wegen $\lambda_1 \neq 0$ folgt daraus $F(W) \subseteq W$. Die eingeschränkte Abbildung $G = F|_W : W \longrightarrow W$ ist daher eine unitäre Abbildung im $(n-1)$-dimensionalen [2] Vektorraum $W$. Nach Induktionsannahme gibt es eine Orthonormalbasis $\mathcal{B}_W = \{v_2, ..., v_n\}$ von $W$ bestehend aus Eigenvektoren von $G$. Nach Definition von $W$ gilt $\langle v_1, v_k\rangle = 0$ für alle $k \in \{2, ..., n\}$ und damit ist $\mathcal{B} = \{v_1, ..., v_n\}$ eine Orthonormalbasis von $V$ aus Eigenvektoren von $F$. □

**Rezept 12.3.8** (zur Bestimmung einer Orthonormalbasis aus Eigenvektoren eines unitären Endomorphismus)**.** Ist $V$ ein $n$-dimensionaler unitärer Vektorraum und $F : V \to V$ unitär, so

- bestimme das charakteristische Polynom

  $$p_F = (\alpha_1 - X)^{a_1} \cdots (\alpha_m - X)^{a_m}$$

  von $F$ – hier ist $|\alpha_1|, ..., |\alpha_m| = 1$, $a_1, ..., a_m \in \mathbb{N}$ mit $a_1 + ...a_m = n$.

- berechne für alle Eigenwerte $\alpha_1, ..., \alpha_m$ die Eigenräume $\mathrm{Eig}(F, \alpha_1), ..., \mathrm{Eig}(F, \alpha_m)$
- damit bekommt man für jeden Eigenraum $\mathrm{Eig}(F, \alpha_k)$ eine Basis $\mathcal{A}_k$, $k = 1, ..., m$.
- wende das Schmidtsche Orthonormalisierungsverfahren auf die Basis $\mathcal{A}_k$ von $\mathrm{Eig}(F, \alpha_k)$ an und erhalte eine Orthonormalbasis $\mathcal{B}_k$ von $\mathrm{Eig}(F, \alpha_k)$
- die Vereinigung $\mathcal{B} = \bigcup_{k=1}^{m} \mathcal{B}_k$ ist dann eine Orthonormalbasis von $V$ bestehend aus Eigenvektoren von $F$.

**Bemerkung 12.3.9.** Das charakteristische Polynom $p_A$ einer orthogonalen Matrix $A \in M_n(\mathbb{R})$ zerfällt nicht immer über $\mathbb{R}$ in Linearfaktoren. Wie man am Beweis von Satz 12.3.7 sieht, gibt es genau dann eine Orthonormalbasis von $\mathbb{R}^n$ aus Eigenvektoren von $A$, wenn $p_A$ über $\mathbb{R}$ in Linearfaktoren zerfällt. Das ist aber, wie wir später an Beispielen sehen werden eher die Ausnahme.

**Lemma 12.3.10.** *Für jedes $n \in \mathbb{N}$ bilden die Mengen*

$$\begin{aligned} U(n) &:= \{A \in GL(n, \mathbb{C}) : A \textit{ ist unitär}\} \\ SU(n) &:= \{A \in U(n) : \det(A) = 1\} \end{aligned}$$

*in der Gruppe $GL(n, \mathbb{C})$ und*

$$\begin{aligned} O(n) &:= \{A \in GL(n, \mathbb{R}) : A \textit{ ist orthogonal}\} \\ SO(n) &:= \{A \in O(n) : \det(A) = 1\} \end{aligned}$$

*in der Gruppe $GL(n, \mathbb{R})$ jeweils Untergruppen.*

[2] vgl. Beweis von Satz 12.2.7

*Beweis.* Wegen $(AB)^* = B^*A^*$ und $(A^*)^* = A$ gilt für orthogonale bzw. unitäre $n \times n$ Matrizen $A, B$ auch $(AB)^{-1} = (AB)^*$ und $(A^{-1})^{-1} = (A^{-1})^*$, dh. $O(n)$ ist eine Untergruppe von $GL(n, \mathbb{R})$ und $U(n)$ ist eine Untergruppe von $GL(n, \mathbb{C})$. Wie bei Korollar 10.2.3 zeigt man $\det(A^*) = \overline{\det(A)}$, damit folgt aus $A^*A = E_n$ und dem Multiplikationssatz für Determinanten $1 = \det E_n = \det(A^*A) = \det(A^*)\det(A) = \overline{\det(A)}\det(A) = |\det(A)|^2$, dh. $\det(A) \in \{\pm 1\}$ für alle $A \in O(n)$ und $\det(A) \in \{z \in \mathbb{C} : |z| = 1\}$ für alle $A \in U(n)$. Nach dem Multiplikationssatz für Determinanten sind also

$$\begin{array}{rcl} \det : O(n) & \to & \{\pm 1\} \\ A & \mapsto & \det(A) \end{array} \qquad \text{und} \qquad \begin{array}{rcl} \det : U(n) & \to & \{z \in \mathbb{C} : |z| = 1\} \\ A & \mapsto & \det(A) \end{array}$$

Gruppenhomomorphismen und daher $SO(n) = \det^{-1}(\{1\})$ und $SU(n) = \det^{-1}(\{1\})$ nach Lemma 3.1.10 Untergruppen von $O(n)$ bzw. $U(n)$. □

**Beispiel 12.3.11.** Betrachtet man die Menge $O(2)$ der orthogonalen $2 \times 2$ Matrizen, so haben diese die Form einer Drehung

$$A_D = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

oder einer Spiegelung

$$A_S = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

mit einem eindeutig bestimmten $\alpha \in [0, 2\pi[$. Denn für jedes $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in O(2)$ folgen wegen der Orthogonalitätsbedingung

$$A^T A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix} = E_2$$

die Gleichungen

$$a^2 + b^2 = 1 \tag{12.3.2}$$
$$c^2 + d^2 = 1 \tag{12.3.3}$$
$$ac + bd = 0 \tag{12.3.4}$$

und wegen (12.3.2) und (12.3.3) lassen sich mit Polardarstellung komplexer Zahlen eindeutig bestimmte $\alpha, \beta \in [0, 2\pi[$ mit

$$a = \cos\alpha, \quad b = \sin\alpha \qquad \text{und} \qquad c = \sin\beta, \quad d = \cos\beta$$

finden. Eingesetzt in (12.3.4) ergibt sich

$$0 = \cos\alpha\sin\beta + \sin\alpha\cos\beta = \sin(\alpha + \beta),$$

was nur für $\alpha + \beta \in \pi\mathbb{Z}$, also $\beta = -\alpha + k\pi$ mit $k \in \mathbb{Z}$ erreicht werden kann und daraus ergeben sich nur zwei Möglichkeiten:

$$c = \sin(\beta) = \sin(-\alpha + k\pi) = \begin{cases} -\sin(\alpha) & \text{für } k \text{ gerade} \\ \sin(\alpha) & \text{für } k \text{ ungerade} \end{cases}$$

$$d = \cos(\beta) = \cos(-\alpha + k\pi) = \begin{cases} \cos(\alpha) & \text{für } k \text{ gerade} \\ -\cos(\alpha) & \text{für } k \text{ ungerade} \end{cases}$$

a) $c = -\sin\alpha$ und $d = \cos\alpha$, dh. $A = A_D = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ ist eine Drehung im Gegenuhrzeigersinn um den Winkel $\alpha$. $p_{A_D} = X^2 - 2X\cos\alpha + 1$ ist das charakteristische Polynom von $A_D$. Deshalb hat $A_D$ nur für $\alpha = 0$ (d.h. $A_D = E_2$) oder $\alpha = \pi$ (d.h. $A_D = -E_2$) Eigenwerte. Wegen $\det(A_D) = 1$ ist $A_D \in SO(2)$.

b) $c = \sin\alpha$ und $d = -\cos\alpha$, dh. $A = A_S = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$ ist eine Spiegelung: $p_{A_S} = -(\cos\alpha + X)(\cos\alpha - X) - (\sin\alpha)^2 = -((\cos\alpha)^2 - X^2) - (\sin\alpha)^2 = X^2 - 1 = (X-1)(X+1)$, es gibt daher Eigenvektoren $\underline{v}_1, \underline{v}_2 \in \mathbb{R}^2$ mit $A\underline{v}_1 = \underline{v}_1$ und $A\underline{v}_2 = -\underline{v}_2$. $\mathbb{R}\underline{v}_1$ ist die Spiegelgerade, $\mathbb{R}\underline{v}_2$ dazu senkrecht. Wegen $\det(A_S) = -1$ ist $A_S \notin SO(2)$.

**Beispiel 12.3.12.** Jede orthogonale Abbildung $F : \mathbb{R}^3 \to \mathbb{R}^3$ mit $\det(F) = 1$ ist durch eine Drehachse und einen Winkel festgelegt: Es sei $A \in SO(3)$ die darstellende Matrix von $F$ in einer Orthonormalbasis, dann folgt durch Einsetzen von $\lambda \neq 0$ in das charakteristische Polynom:

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda E_3) = \det((E_3 - \lambda A^T)A) = \det(E_3 - \lambda A^T) = \det(E_3 - \lambda A) \\ &= (-\lambda)^3 \det(A - \frac{1}{\lambda}E_3) = -\lambda^3 p_A(\frac{1}{\lambda}) \end{aligned}$$

deshalb ergibt $\lambda = 1$ die Gleichung $p_A(1) = -p_A(1)$, dh. $p_A(1) = 0$ und daher ist 1 ein Eigenwert von $F$. Es sei $\underline{v_1} \in \mathbb{R}^3$ ein Eigenvektor von $F$ zum Eigenwert 1 mit $\|\underline{v_1}\| = 1$, dann gilt für $U_1 := \operatorname{lin}\{\underline{v_1}\}$ und $U_2 := \{\underline{w} \in \mathbb{R}^3 : \langle \underline{v_1}, \underline{w}\rangle_{\mathbb{R}^3} = 0\}$ sowohl $F(U_1) \subseteq U_1$ als auch $F(U_2) \subseteq U_2$, denn wegen $F(\underline{v_1}) = \underline{v_1}$ folgt $\langle \underline{v_1}, F(\underline{w})\rangle = \langle F(\underline{v_1}), F(\underline{w})\rangle = \langle \underline{v_1}, \underline{w}\rangle = 0$ für jedes $\underline{w} \in U_2$, dh. $F(\underline{w}) \in U_2$. $F|_{U_2}$ ist somit wieder orthogonal. Bezüglich einer Orthonormalbasis $\mathcal{B} = \{\underline{v_1}\} \cup \mathcal{A}$, die $\underline{v_1}$ als ersten Basisvektor enthält hat $F$ die Blockmatrix $M^{\mathcal{B}}_{\mathcal{B}}(F) = \left(\begin{array}{c|cc} 1 & b_{12} & b_{13} \\ \hline 0 & \multicolumn{2}{c}{M^{\mathcal{A}}_{\mathcal{A}}(F|_{U_2})} \end{array}\right)$ als darstellende Matrix [3], daher ist $\det(M^{\mathcal{A}}_{\mathcal{A}}(F|_{U_2})) = 1$. $\underline{v_1}$ liegt auf der Drehachse und nach Beispiel 12.3.11 gibt es einen eindeutigen Drehwinkel $\alpha \in [0, 2\pi[$ mit $M^{\mathcal{B}}_{\mathcal{B}}(F) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$.

**Satz 12.3.13** (Eulerwinkel)**.** *Zu jedem $A \in SO(3)$ gibt es $\varphi, \psi \in [0, 2\pi[$ und $\theta \in [0, \pi]$, so daß*

$$A = R_3(\psi)R_1(\theta)R_3(\varphi) \tag{12.3.5}$$

*die Hintereinanderausführung der Rotationen*

$$R_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \quad \textit{und} \quad R_3(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

*ist. $\psi, \theta, \varphi$ heißen die* **Eulerwinkel**.

[3] da auch $M^{\mathcal{A}}_{\mathcal{A}}(F|_{U_2})$ orthogonal ist, sind $b_{12} = b_{13} = 0$, denn sowohl in der zweiten und dritten Spalte von $M^{\mathcal{B}}_{\mathcal{B}}(F)$ als auch in den beiden Spalten von $M^{\mathcal{A}}_{\mathcal{A}}(F|_{U_2})$ befinden sich orthogonale Einheitsvektoren

*Beweis.* Es sei $\underline{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \in \mathbb{R}^3$ mit $\|\underline{v}\| = 1$ und $\underline{v} \neq \begin{pmatrix} 0 \\ 0 \\ \pm 1 \end{pmatrix}$, dann gibt es wegen $\|\underline{v}\|^2 = (v_1^2 + v_2^2) + v_3^2 = 1$ ein $\theta \in ]0, \pi[$ mit

$$v_3 = \cos\theta, \quad \sqrt{v_1^2 + v_2^2} = \sin\theta.$$

Da $\sin\theta \neq 0$ ist, gilt $\dfrac{v_1^2}{(\sin\theta)^2} + \dfrac{v_2^2}{(\sin\theta)^2} = 1$, daher gibt es $\psi \in [0, 2\pi[$ mit

$$v_1 = \sin\psi\sin\theta \quad \text{und} \quad v_2 = -\cos\psi\sin\theta,$$

also ist

$$\begin{aligned} R_3(\psi)R_1(\theta)\underline{e}_3 &= \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -\sin\theta \\ \cos\theta \end{pmatrix} = \begin{pmatrix} \sin\psi\sin\theta \\ -\cos\psi\sin\theta \\ \cos\theta \end{pmatrix} = \underline{v} \end{aligned}$$

Ist $\underline{v} = \begin{pmatrix} 0 \\ 0 \\ \pm 1 \end{pmatrix}$, so gilt $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = R_3(0)R_1(0) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} = R_3(0)R_1(\pi) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.
Zu $A \in SO(3)$ setze $\underline{v} := A\underline{e}_3$, dann ist $\|\underline{v}\| = 1$, also gibt es $\theta \in [0, \pi]$ und $\psi \in [0, 2\pi[$ mit $\underline{v} = A\underline{e}_3 = R_3(\psi)R_1(\theta)\underline{e}_3$, woraus

$$\Big[R_3(\psi)R_1(\theta)\Big]^{-1} A\underline{e}_3 = \underline{e}_3 \tag{12.3.6}$$

folgt. Daher ist $\underline{e}_3$ Eigenvektor von $\Big[R_3(\psi)R_1(\theta)\Big]^{-1} A \in SO(3)$ zum Eigenwert 1, daher ist $\Big[R_3(\psi)R_1(\theta)\Big]^{-1} A$ nach Beispiel 12.3.12 eine Rotation um die $\underline{e}_3$-Achse, dh. es gibt ein $\varphi \in [0, 2\pi[$ mit $\Big[R_3(\psi)R_1(\theta)\Big]^{-1} A = R_3(\varphi)$ oder $A = R_3(\psi)R_1(\theta)R_3(\varphi)$. □

**Bemerkung 12.3.14.** Die Eulerwinkel kann man zB. benutzen, um die Orientierung eines starren Körpers mit einem Fixpunkt –etwa ein Kreisel – einmal in einem körperfestes Koordinatensystem und das andere Mal in einem „Laborsystem“ zu beschreiben: Sind beide Koordinatensysteme Orthonormalbasen, so ergibt sich wegen des einen festen Punkts eine Drehung zwischen den beiden Sytemen...

## 12.4 Simultane Diagonalisierung

**Satz 12.4.1.** *Sind $F_1 : V \longrightarrow V, \ldots, F_r : V \longrightarrow V$ unitäre oder selbstadjungierte $\mathbb{K}$-lineare Abbildungen eines unitären Vektorraums $V$ mit $\dim_{\mathbb{K}} V = n < \infty$, so sind äquivalent:*

1. *Es gibt eine Orthonormalbasis $\mathcal{B} = (v_1, \ldots, v_n)$ von $V$, so daß $v_1, \ldots, v_n$ Eigenvektoren aller Abbildungen $F_1, \ldots, F_r$ sind.*

2. *Es gilt $F_iF_j = F_jF_i$ für alle $i, j = 1, ..., r$.*

*Beweis:* Es sei $\mathcal{B} = (v_1, ..., v_n)$ eine Basis wie in (1). Dann gibt es zu jedem $v_k$ eindeutig bestimmte Eigenwerte $\lambda_{l,k} \in \mathbb{K}$ mit $F_l v_k = \lambda_{l,k} v_k$ für alle $l = 1, ..., r$. Damit ist

$$F_j(F_i(v_k)) = F_j(\lambda_{i,k}v_k) = \lambda_{j,k}\lambda_{i,k}v_k = \lambda_{i,k}\lambda_{j,k}v_k = F_i(\lambda_{j,k}v_k) = F_i(F_j(v_k))$$

Entwickelt man nun ein beliebiges Element $v \in V$ in der Basis $\mathcal{B}$, schreibt man also $v = a_1v_1 + ... + a_nv_n$ mit eindeutig bestimmten Koeffizienten $a_1, ..., a_n \in \mathbb{K}$, so ist

$$F_j(F_i(v)) = \sum_{k=1}^{n} a_k F_j(F_i(v_k)) = \sum_{k=1}^{n} a_k F_i(F_j(v_k)) = F_i(F_j(v)).$$

Für die umgekehrte Richtung (2) $\Rightarrow$ (1) betrachte ohne Beschränkung der Allgemeinheit den Fall $r = 2$. Es sei $V = \mathrm{Eig}(F_1, \lambda_1) \oplus ... \oplus \mathrm{Eig}(F_1, \lambda_k)$ die Zerlegung von $V$ in die orthogonalen Eigenräume von $F_1$. Dann ist $F_2(\mathrm{Eig}(F_1, \lambda_j)) \subseteq \mathrm{Eig}(F_1, \lambda_j)$ für $j = 1, ..., k$, denn nach Bedingung (2) gilt für jedes $v_j \in \mathrm{Eig}(F_1, \lambda_j)$:

$$F_1(F_2(v_j)) = F_2(F_1(v_j)) = F_2(\lambda_j v_j) = \lambda_j F_2(v_j).$$

Die Einschränkung von $F_2$ auf jeden der Eigenräume $\mathrm{Eig}(F_1, \lambda_j)$ ist somit wieder unitär oder selbstadjungiert, also gibt es eine Orthonormalbasis von $\mathrm{Eig}(F_1, \lambda_j)$ aus Eigenvektoren von $F_2$. □

## 12.5 Hauptachsentransformation

**Lemma 12.5.1.** *Es sei $A \in M_n(\mathbb{C})$, dann sind äquivalent:*

*a) $A$ ist selbstadjungiert.*

*b) Für jedes $\underline{v} \in \mathbb{C}^n$ ist $\langle \underline{v}, A\underline{v} \rangle_{\mathbb{C}^n} \in \mathbb{R}$.*

*c) Die Sesquilinearform* $\begin{array}{rcl} s_A : \mathbb{C}^n \times \mathbb{C}^n & \to & \mathbb{C} \\ (\underline{x}, \underline{y}) & \mapsto & \langle \underline{x}, A\underline{y} \rangle_{\mathbb{C}^n} \end{array}$ *ist hermitesch.*

*Beweis.*

a)$\Rightarrow$b) Ist $A = A^*$ und $\underline{v} \in \mathbb{C}^n$, so gilt

$$\overline{\langle A\underline{v}, \underline{v} \rangle} = \langle \underline{v}, A\underline{v} \rangle = \langle A^*\underline{v}, \underline{v} \rangle = \langle A\underline{v}, \underline{v} \rangle,$$

dh. $\langle A\underline{v}, \underline{v} \rangle \in \mathbb{R}$.

b)$\Rightarrow$a) Ist $\langle \underline{v}, A\underline{v} \rangle \in \mathbb{R}$ für jedes $\underline{v} \in \mathbb{C}^n$, dann gilt $\langle A^*\underline{v}, \underline{v} \rangle = \langle \underline{v}, A\underline{v} \rangle = \overline{\langle A\underline{v}, \underline{v} \rangle} = \langle A\underline{v}, \underline{v} \rangle$, dh. $A^*\underline{v} = A\underline{v}$ für alle $\underline{v} \in \mathbb{C}^n$, also $A^* = A$.

c)$\Rightarrow$b) Ist $s_A$ hermitesch, dann ist $\langle \underline{x}, A\underline{x} \rangle_{\mathbb{C}^n} = s_A(\underline{x}, \underline{x}) = \overline{s_A(\underline{x}, \underline{x})} = \overline{\langle \underline{x}, A\underline{x} \rangle_{\mathbb{C}^n}}$, also $\langle \underline{x}, A\underline{x} \rangle_{\mathbb{C}^n} \in \mathbb{R}$.

a)⇒c) Ist $A$ selbstadjungiert, so gilt

$$s_A(\underline{x}, \underline{y}) = \langle \underline{x}, A\underline{y} \rangle = \langle A^* \underline{x}, \underline{y} \rangle = \langle A\underline{x}, \underline{y} \rangle = \overline{\langle \underline{y}, A\underline{x} \rangle} = \overline{s_A(y, x)}$$

dh. $s_A$ ist hermitesch. □

**Satz 12.5.2.** *Es sei $s : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ eine Sesquilinearform, so daß $s(\underline{x}, \underline{x}) \in \mathbb{R}$ für jedes $\underline{x} \in \mathbb{C}^n$ ist. Dann gibt es eine selbstadjungierte Matrix $A \in M_n(\mathbb{C})$ und eine Orthonormalbasis $\mathcal{B} = (\underline{v}_1, ..., \underline{v}_n)$ von $\mathbb{C}^n$ bestehend aus Eigenvektoren von $A$ also $A\underline{v}_k = \lambda_k \underline{v}_k$, $\lambda_k \in \mathbb{R}$ so daß für den Koordinatenvektor $\underline{x}' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} \in \mathbb{C}^n$ von $\underline{x} \in \mathbb{C}^n$ bezüglich $\mathcal{B}$, also $\underline{x} = x'_1 \underline{v}_1 + ... + x'_n \underline{v}_n$ gilt:*

$$s(\underline{x}, \underline{y}) = s_A(\underline{x}, \underline{y}) = \sum_{k=1}^{n} \overline{x'_k} \lambda_k y'_k. \tag{12.5.1}$$

*Beweis.* Wie üblich schreiben wir $\underline{x} \in \mathbb{C}^n$ in der Standardbasis $\mathcal{K} = (\underline{e}_1, ..., \underline{e}_n)$. Dann folgt aus der Sesquilinearität von $s$:

$$s(\underline{x}, \underline{y}) = s(\sum_{k=1}^{n} x_k \underline{e}_k, \sum_{l=1}^{n} y_l \underline{e}_l) = \sum_{k,l=1}^{n} \overline{x_k} s(\underline{e}_k, \underline{e}_l) y_l = s_A(\underline{x}, \underline{y})$$

für $A = \begin{pmatrix} s(\underline{e}_1, \underline{e}_1) & \cdots & s(\underline{e}_1, \underline{e}_n) \\ \vdots & & \vdots \\ s(\underline{e}_n, \underline{e}_1) & \cdots & s(\underline{e}_n, \underline{e}_n) \end{pmatrix}$. Wegen $s(\underline{x}, \underline{x}) = \langle \underline{x}, A\underline{x} \rangle_{\mathbb{C}^n} \in \mathbb{R}$ ist $A$ nach Lemma 12.5.1 selbstadjungiert, daher gibt es eine Orthonormalbasis $\mathcal{B} = (\underline{v}_1, ..., \underline{v}_n)$ aus Eigenvektoren von $A$, also $A\underline{v}_k = \lambda_k \underline{v}_k$. Ist also $\underline{x}' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ der Koordinatenvektor von $\underline{x}$ bezüglich $\mathcal{B}$, also $\underline{x} = \sum_{k=1}^{n} x'_k \underline{v}_k$, so ist

$$\langle \underline{x}, A\underline{y} \rangle_{\mathbb{C}^n} = \sum_{k,l=1}^{n} \overline{x'_k} \langle \underline{v}_k, A\underline{v}_l \rangle_{\mathbb{C}^n} y'_l = \sum_{k=1}^{n} \overline{x'_k} \lambda_k y'_k.$$

□

**Lemma 12.5.3.** *Es sei $B \in M_n(\mathbb{R})$ und $q_B$ die quadratische Form* $\begin{array}{rcl} q_B : \mathbb{R}^n & \to & \mathbb{R} \\ \underline{x} & \mapsto & \langle \underline{x}, B\underline{x} \rangle \end{array}$ *zu $B$. Dann gibt es eine selbstadjungierte Matrix $A \in M_n(\mathbb{R})$ mit $q_A = q_B$.*

*Beweis.* Es sei $B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} \in M_n(\mathbb{R})$ und $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, dann ist

$$q_B(\underline{x}) = \sum_{k,l=1}^{n} x_k x_l b_{kl} = \sum_{k,l=1}^{n} x_k x_l (\frac{1}{2} b_{kl} + \frac{1}{2} b_{lk}) = q_A(\underline{x})$$

für $A = \begin{pmatrix} b_{11} & \frac{1}{2}(b_{12}+b_{21}) & & \cdots & \frac{1}{2}(b_{1n}+b_{n1}) \\ \frac{1}{2}(b_{21}+b_{12}) & b_{22} & & & \vdots \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ \frac{1}{2}(b_{n1}+b_{1n}) & & & & b_{nn} \end{pmatrix}$, was offenbar selbstadjungiert ist. □

**Satz 12.5.4.** *Ist $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine Bilinearform und $q : \mathbb{R}^n \to \mathbb{R}$ die zugehörige quadratische Form, dann gibt es eine selbstadjungierte Matrix $A \in M_n(\mathbb{R})$ mit $q = q_A$.*

*Beweis.* Verwende $B = \begin{pmatrix} b(\underline{e}_1, \underline{e}_1) & \cdots & b(\underline{e}_1, \underline{e}_n) \\ \vdots & & \vdots \\ b(\underline{e}_n, \underline{e}_1) & \cdots & b(\underline{e}_n, \underline{e}_n) \end{pmatrix}$, dann gilt wie eben $q = q_B$ und nach Lemma 12.5.3 gilt auch noch $q_B = q_A$ für

$$A = \begin{pmatrix} b(\underline{e}_1, \underline{e}_1) & \frac{1}{2}(b(\underline{e}_1, \underline{e}_2) + b(\underline{e}_2, \underline{e}_1)) & & \cdots & \frac{1}{2}(b(\underline{e}_1, \underline{e}_n) + b(\underline{e}_n, \underline{e}_1)) \\ \frac{1}{2}(b(\underline{e}_1, \underline{e}_2) + b(\underline{e}_2, \underline{e}_1)) & b(\underline{e}_2, \underline{e}_2) & & & \vdots \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ \frac{1}{2}(b(\underline{e}_1, \underline{e}_n) + b(\underline{e}_n, \underline{e}_1)) & & & & b(\underline{e}_n, \underline{e}_n) \end{pmatrix}.$$

□

**Bemerkung 12.5.5.** Viele physikalische Größen – wie etwa kinetische Energie $\frac{1}{2}mv^2$ – sind in eindimensionalen Modellen quadratische Polynome und daraus werden in zwei- bzw. dreidimensionalen Modellen quadratische Formen. Die eben aufgeführten Aussagen, erlauben durch Wahl von geeigneten Orthonormalbasen die Berechnung dieser quadratischen Formen oft radikal zu vereinfachen.

# Kapitel 13

# Topologie

## 13.1 Topologische Räume und stetige Abbildungen

**Definition 13.1.1.** *Es sei $X \neq \emptyset$ eine Menge und $\mathcal{O} \subseteq \mathcal{P}(X) := \{Y \subseteq X\}$ ein System von Teilmengen mit*

*a)* $\emptyset, X \in \mathcal{O}$.

*b) Ist $I$ eine Menge und sind für $i \in I$ Elemente $V_i \in \mathcal{O}$ gegeben, so gilt:*

$$\bigcup_{i \in I} V_i \in \mathcal{O} \tag{13.1.1}$$

*c) Ist $J$ eine endliche Menge und sind für $i \in J$ Elemente $V_i \in \mathcal{O}$ gegeben, so gilt:*

$$\bigcap_{i \in J} V_i \in \mathcal{O}, \tag{13.1.2}$$

*dann heißt $\mathcal{O}$ eine* **Topologie** *auf $X$ und $(X, \mathcal{O})$ ein* **topologischer Raum**. *Es sei $(X, \mathcal{O})$ ein topologischer Raum, dann*

- *heißen die Elemente $V \in \mathcal{O}$* **offene Mengen**.
- *Ist $x \in X$, so heißt $U \subseteq X$ eine* **Umgebung** *von $x$, wenn es $V \in \mathcal{O}$ gibt mit $x \in V \subseteq U$.*
- *Ist $x \in X$, so heißt $\mathcal{U}_x \subseteq \mathcal{P}(X)$ eine* **Umgebungsbasis** *von $x$, wenn es für jede Umgebung $U$ von $x$ eine Umgebung $V$ von $x$ mit $V \in \mathcal{U}_x$ und $V \subseteq U$ gibt.*
- *$(X, \mathcal{O})$ heißt* **hausdorffsch**, *wenn es für alle $x, y \in X$ mit $x \neq y$ offene Mengen $U, V \in \mathcal{O}$ gibt mit $U \cap V = \emptyset$ und $x \in U$ und $y \in V$.*
- *Eine Menge $A \subseteq X$ heißt* **abgeschlossen**, *wenn es $V \in \mathcal{O}$ gibt mit $A = X \backslash V$.*
- *$\mathcal{B} \subseteq \mathcal{O}$ heißt eine* **Basis der Topologie** *$\mathcal{O}$, wenn sich jedes $U \in \mathcal{O}$ als Vereinigung $U = \bigcup_{i \in I} U_i$ von offenen Mengen $U_i \in \mathcal{B}$ schreibt.*

*Ist $X$ eine Menge und sind $\mathcal{O}$ und $\tilde{\mathcal{O}}$ Topologien auf $X$ mit $\mathcal{O} \subseteq \tilde{\mathcal{O}}$, dann heißt $\mathcal{O}$* **gröber** *als $\tilde{\mathcal{O}}$ (oder $\tilde{\mathcal{O}}$* **feiner** *als $\mathcal{O}$). Ist $(X, \mathcal{O})$ ein topologischer Raum, $Y \subseteq X$, so heißt $x \in X$*

- **innerer Punkt** *von $Y$, wenn $Y$ eine Umgebung von $x$ ist*
- **äußerer Punkt** *von $Y$, wenn $x$ ein innerer Punkt von $X \backslash Y$ ist.*
- **isolierter Punkt** *von $Y$, wenn es eine Umgebung $U$ von $x$ gibt mit $U \cap Y = \{x\}$*
- **Berührungspunkt** *von $Y$, wenn für jede Umgebung $U$ von $x$ gilt: $Y \cap U \neq \emptyset$.*
- **Häufungspunkt** *von $Y$, wenn für jede Umgebung $U$ von $x$ gilt: $U \cap (Y \backslash \{x\}) \neq \emptyset$.*

$$\overline{Y} := \{x \in X : \text{ für jede Umgebung } V \text{ von } x \text{ gilt: } V \cap Y \neq \emptyset\} \tag{13.1.3}$$

*heißt der* **Abschluß** *von $Y$.*

$$\overset{\circ}{Y} := \{x \in Y : x \text{ ist innerer Punkt von } Y\} \tag{13.1.4}$$

*heißt* **offener Kern** *von $Y$.*

$$\partial Y := \overline{Y} \cap \overline{X \backslash Y} \tag{13.1.5}$$

*heißt der* **Rand** *von $Y$. $Y \subseteq X$ heißt* **dicht***, wenn $\overline{Y} = X$ ist. $(X, \mathcal{O})$ ist ein* **separabler topologischer Raum***, wenn es eine abzählbare dichte Teilmenge gibt.*

**Beispiel 13.1.2.**

a) Ist $X$ eine Menge, so bilden $\mathcal{O} = \{\emptyset, X\}$ und $\mathcal{O} = \mathcal{P}(X)$ immer eine Topologie auf $X$.

b) Es sei $X = \{\square, \diamond\}$, dann erfüllt $\mathcal{O} = \{\emptyset, \{\square\}, X\}$, wie man durch Bilden aller möglichen Durchschnitte und Vereinigungen nachprüfen kann, alle Eigenschaften einer Topologie. Diese ist jedoch nicht hausdorffsch, denn $\{\square\}$ und $X$ sind alle Umgebungen von $\square$ und $X$ ist die einzige Umgebung von $\diamond$, daher gibt es keine disjunkten Umgebungen von $\square$ und $\diamond$. In $\mathcal{O}$ sind $\emptyset = X \backslash X$, $\{\diamond\} = X \backslash \{\square\}$ und $X = X \backslash \emptyset$ alle abgeschlossenen Mengen.

c) Jeder metrische Raum $(X, d)$ wird durch die Wahl von

$$\mathcal{O}_d := \Big\{ U \subseteq X \ : \ \text{für alle } x \in U \text{ existiert } \varepsilon = \varepsilon(x, U) > 0 \text{ mit } K(x, \varepsilon) := \{y \in X : d(x, y) < \varepsilon\} \subseteq U \Big\} \tag{13.1.6}$$

zu einem topologischen hausdorffschen Raum $(X, \mathcal{O}_d)$. Zu jedem $a \in X$ und $r > 0$ ist die offene Kugel $K(a, r) := \{x \in X : d(x, a) < r\}$ offen bzgl. $\mathcal{O}_d$ und damit hat jeder Punkt $a \in X$ die Umgebungsbasis $\{K(a, r) : r > 0\}$ und die abzählbare Umgebungsbasis $\{K(a, \frac{1}{n}) : n \in \mathbb{N}\}$.

*Beweis.* • $\emptyset, X \in \mathcal{O}_d$

- Ist $I$ eine Menge und $U_i \in \mathcal{O}_d$ für alle $i \in I$, $w \in \bigcup_{i\in I} U_i$, so wähle $j \in I$ mit $w \in U_j$ und $r := r(w, U_j)$, dann ist $\{v \in X : d(v,w) < r(w,U_j)\} \subseteq U_j \subseteq \bigcup_{i\in I} U_i$, also $\bigcup_{i\in I} U_i \in \mathcal{O}_d$.
- Ist $J$ eine endliche Menge und $U_j \in \mathcal{O}_d$ für alle $j \in J$ und $w \in \bigcap_{i\in J} U_i$, dann ist $r := \min\{r(w,U_j) : j \in J\} > 0$ und
$$\{v \in X : d(v,w) < r\} \subseteq \{v \in X : d(v,w) < r(w,U_j)\} \subseteq U_j$$
für alle $j \in J$, also $\{v \in X : d(v,w) < r\} \subseteq \bigcap_{j\in J} U_j$, dh. $\bigcap_{j\in J} U_j \in \mathcal{O}_d$

und somit ist $\mathcal{O}_d$ eine Topologie auf $X$. Ist $a \in X$ und $r > 0$, dann ist die offene Kugel $K(a,r) := \{x \in X : d(x,a) < r\}$ auch offen in der von $d$ erzeugten Topologie $\mathcal{O}_d$ auf $X$. Denn für jedes $x \in K(a,r)$ ist $\delta(x) := \frac{1}{2}(r - d(x,a)) > 0$, also folgt aus der Dreiecksungleichung $d(y,a) \leq d(y,x) + d(x,a) < \delta(x) + d(x-a) < r$ für jedes $y \in X$ mit $d(x,y) < \delta(x)$, dh. $K(x,\delta(x)) \subseteq K(a,r)$ für jedes $x \in K(a,r)$. Dies zeigt $K(a,r) \in \mathcal{O}_d$. Zu $a \in X$ bilden also die **offenen Kugeln**

$$K(a,r) := \{y \in X : d(x,y) < r\} \quad \text{für } r > 0$$

und ebenso $K(a, \frac{1}{n}), n \in \mathbb{N}$ eine Umgebungsbasis von $a$, denn nach Definition von $\mathcal{O}_d$ gibt es für jede offene Umgebung $U$ von $a$ ein $r > 0$ mit $K(a,r) \subseteq U$ und $K(a,r) \in \mathcal{O}_d$. □

d) Ist $(V, \|\cdot\|)$ ein normierter Vektorraum, dann wird durch

$$\begin{aligned} \mathcal{O}_{\|\cdot\|} \;:=\; \{U \subseteq V :\ & \text{Für jedes } u \in U \text{ gibt es ein } r = r(u,U) > 0 \text{ mit} \\ & \{w \in V : \|u - w\| < r\} \subseteq U\} \end{aligned}$$

eine Topologie auf $V$ definiert – diese heißt die **Normtopologie** auf $V$. Dies ist ein Spezialfall von (c).

e) Auf $\mathbb{R}$ und $\mathbb{C}$ wird durch den Absolutbetrag $|\cdot|$ nach (d) eine Normtopologie definiert. Diese wird als **Standardtopologie** bezeichnet und sofern nicht explizit anders vermerkt betrachten wir $\mathbb{R}$ oder $\mathbb{C}$ immer als topologischen Raum versehen mit dieser Standardtopologie.

**Lemma 13.1.3.** *Es sei $(X, \mathcal{O})$ ein topologischer Raum und $\emptyset \neq Y \subseteq X$. Dann definiert*

$$\mathcal{O}_Y := \{Y \cap V : V \in \mathcal{O}\} \tag{13.1.7}$$

*eine Topologie auf $Y$, die sogenannte* **Relativtopologie**.

*Beweis.* $\emptyset = \emptyset \cap Y, Y = Y \cap X \in \mathcal{O}_Y$. Ist $I$ eine Menge und $V_i \in \mathcal{O}_Y$ für alle $i \in I$, so gibt es $W_i \in \mathcal{O}$ mit $V_i = Y \cap W_i$, daher ist $\bigcup_{i\in I} W_i \in \mathcal{O}$ und daher

$$\bigcup_{i\in I} V_i = \bigcup_{i\in I}(Y \cap W_i) = Y \cap \left(\bigcup_{i\in I} W_i\right) \in \mathcal{O}_Y.$$

Ist $J$ eine endliche Menge, $V_j \in \mathcal{O}_Y$ für alle $j \in J$, also ist $V_j = Y \cap W_j$ mit $W_j \in \mathcal{O}$, dann ist

$$\bigcap_{i\in J} V_i = \bigcap_{i\in J}(Y \cap W_i) = Y \cap \left(\bigcap_{i\in J} W_i\right) \in \mathcal{O}_Y.$$

□

**Lemma 13.1.4.** *Es sei $\emptyset \neq X$ eine Menge und für $i \in I$ sei $\mathcal{O}_i \subseteq \mathcal{P}(X)$ eine Topologie auf $X$, dann ist auch*

$$\mathcal{O} := \bigcap_{i\in I} \mathcal{O}_i = \{V \subseteq X : V \in \mathcal{O}_i \text{ für alle } i \in I\} \tag{13.1.8}$$

*eine Topologie auf $X$. Insbesondere ist für jede Teilmenge $\mathcal{S} \subseteq \mathcal{P}(X)$ das System*

$$\mathcal{O}_{\mathcal{S}} := \bigcap_{\substack{\mathcal{O} \text{ Topologie auf } X \\ \mathcal{S}\subseteq\mathcal{O}}} \mathcal{O} \tag{13.1.9}$$

*eine Topologie auf $X$, nämlich die gröbste Topologie auf $X$, in der alle Mengen aus $\mathcal{S}$ offen sind.*

*Beweis.* Sind $\mathcal{O}_i$ Topologien auf $X$ für alle $i \in I$, dann ist $\emptyset, X \in \mathcal{O}_i$ für alle $i \in I$, also $\emptyset, X \in \mathcal{O}$. Ist $J$ eine Menge und $V_j \in \mathcal{O}$ für alle $j \in J$, dann ist $V_j \in \mathcal{O}_i$ für alle $i \in I$, also $\bigcup_{j\in J} V_j \in \mathcal{O}_i$ für alle $i \in I$, dh. $\bigcup_{j\in J} V_j \in \mathcal{O}$. Ist $J$ eine endliche Menge und $V_j \in \mathcal{O}$ für alle $j \in J$, dann ist $V_j \in \mathcal{O}_i$ für alle $i \in I$, also $\bigcap_{j\in J} V_j \in \mathcal{O}_i$ für alle $i \in I$, dh. $\bigcap_{j\in J} V_j \in \mathcal{O}$. □

**Lemma 13.1.5.** *Ist $(X, \mathcal{O})$ ein topologischer Raum, $Y \subseteq X$ dann sind äquivalent:*

*a) $Y$ ist offen*

*b) Jedes $y \in Y$ ist ein innerer Punkt von $Y$, dh. $Y = \overset{\circ}{Y}$.*

*Beweis.*

a)⇒b) Ist $Y$ offen, $y \in Y$, so ist $Y$ eine Umgebung von $y$, also $y$ ein innerer Punkt von $Y$, dh. $Y \subseteq \overset{\circ}{Y}$ und da immer $\overset{\circ}{Y} \subseteq Y$ gilt, folgt daher $\overset{\circ}{Y} = Y$.

b)⇒a) Ist $y \in Y$, so gibt es eine offene Umgebung $U(y)$ von $y$ mit $U(y) \subseteq Y$, daher ist $Y = \bigcup_{y\in Y} U(y)$ offen. □

**Lemma 13.1.6.** *Ist $(X, \mathcal{O})$ ein topologischer Raum, $Y \subseteq X$ dann sind äquivalent:*

*a) $Y$ ist abgeschlossen*

*b) Jeder Berührpunkt $y$ von $Y$ liegt in $Y$, dh. $Y = \overline{Y}$.*

*Beweis.* $Y$ ist genau dann abgeschlossen, wenn $X \backslash Y$ offen ist, was nach Lemma 13.1.5 genau dann der Fall ist, wenn jeder Punkt $z \in X \backslash Y$ ein innerer Punkt von $X \backslash Y$ ist. Das aber bedeutet, daß es für jedes $z \in X \backslash Y$ eine Umgebung $U_z \subseteq X \backslash Y$ von $z$ gibt, weshalb kein $z \in X \backslash Y$ Berührpunkt von $Y$ sein kann. Da jedes $y \in Y$ auch Berührpunkt von $Y$ ist, gilt dann auch $Y = \overline{Y}$. □

**Beispiel 13.1.7.**
Es sei $(\mathbb{C}, \mathcal{O}_{\mathbb{C}})$ die Standardtopologie auf $\mathbb{C}$ wie in Beispiel 13.1.2 und $Y := \{x + iy \in \mathbb{C} : x, y \in [-1, 1[\}$. Dann ist $-1 \in Y$

- kein innerer Punkt von $Y$, denn jede Umgebung von $-1$ enthält nach Definition von $\mathcal{O}$ eine offene Kugel $K(-1, \varepsilon)$ mit $\varepsilon > 0$ und nach Definition von $Y$ liegt keine solche Kugel in $Y$.
- kein isolierter Punkt von $Y$, denn $K(-1, \varepsilon) \cap Y \neq \{-1\}$ für jedes $\varepsilon > 0$.
- Berührpunkt von $Y$, denn für jedes $\varepsilon > 0$ ist $K(-1, \varepsilon) \cap Y \neq \emptyset$.
- Häufungspunkt von $Y$, denn für jedes $\varepsilon > 0$ ist $K(-1, \varepsilon) \cap (Y \backslash \{-1\} \neq \emptyset$.

**Lemma 13.1.8.** *Es sei $(X, \mathcal{O})$ ein topologischer Raum und $Y \subseteq X$, dann ist $\overline{Y}$ die kleinste abgeschlossene Menge, die $Y$ enthält und $\overset{\circ}{Y}$ die größte offene Menge, die in $Y$ enthalten ist. In diesem Fall gilt: $X \backslash \overline{Y} = (X \backslash Y)^\circ$ und $X \backslash \overset{\circ}{Y} = \overline{X \backslash Y}$.*

*Beweis.* Nach Definition ist $x$ genau dann ein innerer Punkt von $Y$, wenn es eine offene Umgebung $U_x$ von $x$ gibt mit $U_x \subseteq Y$. Der offene Kern $\overset{\circ}{Y}$ enthält also gerade alle offenen Mengen, die in $Y$ enthalten sind und ist daher als Vereinigung dieser offenen Mengen wieder offen, also die größte offene Menge, die in $Y$ enthalten ist. $x \in X$ ist nicht im Abschluß $\overline{Y}$ von $Y$ genau dann wenn es eine Umgebung $U_x$ von $x$ mit $U_x \cap Y = \emptyset$ gibt, dh. mit $U_x \subseteq X \backslash Y$, dh es gilt $X \backslash \overline{Y} \subseteq (X \backslash Y)^\circ$. $x \in (X \backslash Y)^\circ$ genau dann wenn es eine Umgebung $U_x$ von $x$ mit $U_x \subseteq X \backslash Y$ gibt, also $U_x \cap Y = \emptyset$, dh. $x$ ist kein Berührpunkt von $Y$ oder $(X \backslash Y)^\circ \subseteq X \backslash \overline{Y}$. Zusammengenommen gilt

$$X \backslash \overline{Y} = (X \backslash Y)^\circ. \tag{13.1.10}$$

Wende (13.1.10) für $X \backslash Y$ statt $Y$ an, so folgt $X \backslash \overline{(X \backslash Y)} = [X \backslash (X \backslash Y)]^\circ = Y^\circ$ und nach Komplementbildung folgt

$$X \backslash (X \backslash \overline{(X \backslash Y)}) = \overline{X \backslash Y} = X \backslash \overset{\circ}{Y}.$$

$\overline{Y} = X \backslash ((X \backslash Y)^\circ)$ ist als Komplement der offenen Menge $(X \backslash Y)^\circ$ abgeschlossen und da $(X \backslash Y)^\circ$ die größte in $X \backslash Y$ enthaltene offene Menge ist, ist $\overline{Y}$ die kleinste abgeschlossene Menge, die $Y$ enthält. □

**Lemma 13.1.9.** *Es sei $(X, \mathcal{O}_X)$ ein topologischer Raum und $Y \subseteq X$, dann sind äquivalent:*

*a) $Y$ ist dicht*

*b) Für jedes $V \in \mathcal{O}_X$, $V \neq \emptyset$ gilt $Y \cap V \neq \emptyset$.*

*Beweis.* Es sei $Y$ dicht in $X$, dh. $\overline{Y} = X$ und $\emptyset \neq V \in \mathcal{O}_X$. Angenommen $Y \cap V = \emptyset$, so ist $V \subseteq X \backslash Y$ und weil $V$ offen ist, folgt daraus sogar $\emptyset \neq V = \overset{\circ}{V} \subseteq (X \backslash Y)^{\circ} = X \backslash \overline{Y} = \emptyset$, ein offensichtlicher Widerspruch. Ist umgekehrt $Y \cap V \neq \emptyset$ für jedes $V \in \mathcal{O}_X$ mit $\emptyset \neq V$ und $x \in X$, dann gibt es für jede Umgebung $U$ von $x$ ein $V \in \mathcal{O}_X$ mit $x \in V$ und $V \subseteq U$, also ist $Y \cap U \supseteq Y \cap V \neq \emptyset$ und somit $x$ ein Berührungspunkt von $Y$, dh. $\overline{Y} = X$. □

**Beispiel 13.1.10.** In $(\mathbb{R}, \mathcal{O}_{\mathbb{R}})$ sind $\mathbb{Q}$ und $\mathbb{R} \backslash \mathbb{Q}$ dichte Mengen, denn zu jedem $V \in \mathcal{O}_{\mathbb{R}}$, $V \neq \emptyset$ wähle $x \in V$. Dann gibt es nach Definition von $\mathcal{O}_{\mathbb{R}}$ ein $r > 0$ mit $]x-r, x+r[ \subseteq V$. Nach Lemma 4.1.3 gibt es ein $q \in ]x-r, x+r[ \cap \mathbb{Q}$, also ist $\mathbb{Q}$ dicht in $(\mathbb{R}, \mathcal{O}_{\mathbb{R}})$. $]x-r, x+r[$ ist überabzählbar, also $]x-r, x+r[ \cap (\mathbb{R} \backslash \mathbb{Q}) \neq \emptyset$, da $\mathbb{Q}$ abzählbar ist.

**Definition 13.1.11.** *Ist $(X, \mathcal{O})$ ein topologischer Raum, $I$ eine gerichtete Menge und $(x_i)_{i \in I}$ ein Netz in $X$, dann heißt $a \in X$ ein*

*a)* **Grenzwert** *von $(x_i)_{i \in I}$, genau dann wenn es für jede Umgebung $U$ von $a$ ein $j(U) \in I$ gibt, so daß $x_i \in U$ für alle $i \in I$ mit $j(U) \leq i$ gilt. Schreibweise: $a = \lim\limits_{i \in I} x_i$.*

*b)* **Häufungswert** *von $(x_i)_{i \in I}$, genau dann wenn für jede Umgebung $U$ von $a$ und für alle $i \in I$ ein $j \in I$ mit $j \geq i$ und $x_j \in U$ gibt.*

**Lemma 13.1.12.** *Ist $(X, \mathcal{O})$ ein hausdorffscher topologischer Raum, dann hat jedes Netz $(x_i)_{i \in I}$ in $X$ höchstens einen Grenzwert.*

*Beweis.* Sind $a, b \in X$ mit $a \neq b$ Grenzwerte von $(x_i)_{i \in I}$, dann gibt es eine Umgebung $U$ von $a$ und eine Umgebung $V$ von $b$ mit $U \cap V = \emptyset$. Nach Definition des Grenzwerts gibt es $j(U), j(V) \in I$, so daß $x_k \in U$ für alle $k \in I$ mit $k \geq j(U)$ ist und $x_l \in V$ für alle $l \in I$ mit $l \geq j(V)$ ist. Da $I$ gerichtet ist, gibt es ein $j \in I$ mit $j \geq j(U)$ und $j \geq j(V)$, also ist $x_j \in U \cap V = \emptyset$, ein Widerspruch. □

**Lemma 13.1.13.** *Es sei $(X, \mathcal{O}_X)$ ein topologischer Raum, $A \subseteq X$ und $a \in X$, dann sind äquivalent:*

*a) $a \in \overline{A}$*

*b) Es gibt ein Netz $(x_i)_{i \in I}$ in $A$ mit $a = \lim\limits_{i \in I} x_i$.*

*Beweis.* Es sei $\mathcal{U}(a)$ die Menge aller Umgebungen von $a$.

a)⇒b) Ist $a \in \overline{A}$, so gilt $U \cap A \neq \emptyset$ für jede Umgebung $U$ von $a$. Wir können also zu $U \in \mathcal{U}(a)$ ein $x_U \in U \cap A$ wählen, dann ist $(x_U)_{U \in \mathcal{U}(a)}$ ein Netz in $A$ mit $a = \lim\limits_{U \in \mathcal{U}(a)} x_U$.

b)⇒a) Ist $(x_i)_{i \in I}$ ein Netz in $A$ mit $a = \lim\limits_{i \in I} x_i$ und $U$ eine Umgebung von $a$, dann gibt es $j(U) \in I$ mit $x_i \in U$ für alle $i \in I, i \geq j(U)$, dh. $\emptyset \neq \{x_i : i \geq j(U)\} \subseteq U \cap A$, also ist $a$ Berührpunkt von $A$.

□

**Definition 13.1.14.** *Es seien $(X, \mathcal{O}_X)$ und $(Y, \mathcal{O}_Y)$ topologische Räume, $a \in X$ und $f : X \to Y$. Dann heißt $f$* **stetig im Punkt** *$a$, wenn für jede Umgebung $V$ von $f(a)$ auch $f^{-1}(V)$ eine Umgebung von $a$ ist. $f$ ist* **stetig**, *wenn $f$ in jedem Punkt $a \in X$ stetig ist. $f : X \to Y$ heißt ein* **Homöomorphismus**, *wenn $f$ stetig und bijektiv ist und auch die Umkehrabbildung $f^{-1} : Y \to X$ stetig ist.*

**Satz 13.1.15.** *Es seien $(X, \mathcal{O}_X)$ und $(Y, \mathcal{O}_Y)$ topologische Räume, $f : X \to Y$. Dann sind äquivalent:*

*a) $f$ ist stetig*

*b) $f^{-1}(V) \in \mathcal{O}_X$ für alle $V \in \mathcal{O}_Y$*

*c) $f(\overline{B}) \subseteq \overline{f(B)}$ für alle $B \subseteq X$*

*d) $f^{-1}(A)$ ist abgeschlossen für alle abgeschlossenen Mengen $A$ in $Y$.*

*Beweis.*

a) ⇒ c) Es sei $B \subseteq X$ und $x \in X$ mit $x \in \overline{B}$ und es sei $V$ eine Umgebung von $f(x)$. Wegen der Stetigkeit von $f$ im Punkt $x$ ist dann $f^{-1}(V)$ eine Umgebung von $x$, also $f^{-1}(V) \cap B \neq \emptyset$ nach der Definition des Abschlußes von $B$. Damit ist

$$\emptyset \neq f(f^{-1}(V) \cap B) = f(\{y \in B : f(y) \in V\}) = V \cap f(B),$$

also $f(x) \in \overline{f(B)}$.

c) ⇒ d) Es sei $A \subseteq Y$ abgeschlossen und $B := f^{-1}(A)$. Nach Voraussetzung in c) folgt: [1]

$$f(\overline{B}) \subseteq \overline{f(B)} = \overline{f(f^{-1}(A))} = \overline{f(X) \cap A} \subseteq \overline{A} = A$$

Durch Urbildbildung folgt:

$$\overline{B} \subseteq f^{-1}(f(\overline{B})) \subseteq f^{-1}(A) = B \subseteq \overline{B}$$

oder $B = \overline{B}$.

d) ⇒ b) Für alle $W \subseteq Y$ ist $X \backslash f^{-1}(W) = f^{-1}(Y \backslash W)$, daher ist $f^{-1}(V) = X \backslash f^{-1}(Y \backslash V)$ für jedes $V \in \mathcal{O}_Y$ das Komplement der abgeschlossenen Menge $f^{-1}(Y \backslash V)$ und damit offen.

b) ⇒ a) Es sei $x \in X$ und $U$ eine Umgebung von $f(x)$, $V \in \mathcal{O}_Y$ mit $V \subseteq U$ und $f(x) \in V$, dann ist nach b) $f^{-1}(V)$ offen; ferner $x \in f^{-1}(V)$, also sind $f^{-1}(V) \subseteq f^{-1}(U)$ Umgebungen von $x$.

□

**Satz 13.1.16.** *Sind $(X, \mathcal{O}_X)$ und $(Y, \mathcal{O}_Y)$ topologische Räume, $a \in X$ und $f : X \to Y$ eine Funktion, dann sind äquivalent:*

[1] wegen $f(f^{-1}(A)) = \{f(x) : x \in X, f(x) \in A\} = f(X) \cap A$

*a) $f$ ist stetig in $a$.*

*b) Für jedes Netz $(x_i)_{i\in I}$ in $X$ mit $a = \lim\limits_{i\in I} x_i$ gilt $f(a) = \lim\limits_{i\in I} f(x_i)$.*

*Beweis.*

a)⇒b) Ist $V$ Umgebung von $f(a)$, dann ist $U := f^{-1}(V)$ eine Umgebung von $a$ mit $f(U) \subseteq V$. Ist nun $(x_i)_{i\in I}$ ein Netz in $X$ mit $a = \lim\limits_{i\in I} x_i$, so gibt es $j(U) \in I$ mit $x_i \in U$ für alle $i \geq j(U)$ und damit ist $f(x_i) \in V$ für alle $i \geq j(U)$, dh. $f(a) = \lim\limits_{i\in I} f(x_i)$.

b)⇒a) Angenommen $f$ ist in $a$ nicht stetig, so gibt es eine Umgebung $V$ von $f(a)$ mit $f(U) \not\subseteq V$ für jede Umgebung $U$ von $a$. Betrachte die Menge $\mathcal{U}(a)$ aller Umgebungen von $a$; diese ist gerichtet. Daher gibt es für alle $U \in \mathcal{U}(a)$ ein $x_U \in U$ mit $f(x_U) \notin V$. Somit definiert $(x_U)_{U\in\mathcal{U}(a)}$ ein Netz in $X$ mit $a = \lim\limits_{U\in\mathcal{U}(a)} x_U$ aber $f(a) \neq \lim\limits_{U\in\mathcal{U}(a)} f(x_U)$.

□

**Satz 13.1.17.** *Es seien $(X, \mathcal{O}_X)$, $(Y, \mathcal{O}_Y)$ und $(Z, \mathcal{O}_Z)$ topologische Räume, $f : X \to Y$ und $g : Y \to Z$. Dann gilt:*

- *Ist $f$ stetig in $a \in X$ und $g$ stetig in $f(a)$, so ist auch $g \circ f$ stetig in $a$.*
- *Sind $f$ und $g$ stetig, so ist auch $g \circ f$ stetig.*

*Beweis.* Ist $W$ eine Umgebung von $(g \circ f)(a)$, so ist wegen der Stetigkeit von $g$ im Punkt $f(a)$ auch $g^{-1}(W)$ eine Umgebung von $f(a)$. Ebenso ist $(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W))$ wegen der Stetigkeit von $f$ in $a$ eine Umgebung von $a$, also ist $g \circ f$ in $a$ stetig. □

**Lemma 13.1.18.** *Es seien $(X, \mathcal{O}_X)$ und $(Y, \mathcal{O}_Y)$ topologische Räume, $f : X \to Y$ eine Abbildung, $a \in X$ und $\mathcal{U}_{f(a)}$ eine Umgebungsbasis von $f(a)$ bezüglich $\mathcal{O}_Y$, $\mathcal{B}$ eine Basis der Topologie $\mathcal{O}_Y$. Dann gilt:*

*a) $f$ ist in $a$ genau dann stetig, wenn $f^{-1}(U)$ für jedes $U \in \mathcal{U}_{f(a)}$ eine Umgebung von $a$ ist.*

*b) $f$ ist genau dann stetig, wenn $f^{-1}(U) \in \mathcal{O}_X$ für alle $U \in \mathcal{B}$ gilt.*

*Beweis.*

a) Nach Definition ist $f$ genau dann in $a$ stetig, wenn für jede Umgebung $V$ von $f(a)$ auch $f^{-1}(V)$ eine Umgebung von $a$ ist. Da jedes $U \in \mathcal{U}_{f(a)}$ eine Umgebung von $f(a)$ ist zeigt dies die Folgerung „⇒“ in a). Nach Definition einer Umgebungsbasis von $f(a)$ gibt es für jede Umgebung $V$ von $f(a)$ eine Umgebung $U \in \mathcal{U}_{f(a)}$ mit $U \subseteq V$. Ist also $f^{-1}(U)$ eine Umgebung von $a$, dann ist $f^{-1}(U) \subseteq f^{-1}(V)$, also auch $f^{-1}(V)$ Umgebung von $a$ und damit $f$ stetig in $a$, was auch die andere Richtung zeigt.

b) Nach Satz 13.1.15 ist $f$ genau dann stetig, wenn $f^{-1}(U) \in \mathcal{O}_X$ für alle $U \in \mathcal{O}_Y$ erfüllt ist. Da $\mathcal{B} \subseteq \mathcal{O}_Y$ gilt, folgt „$\Rightarrow$“; für die andere Implikation „$\Leftarrow$“ beachte, daß jedes $U \in \mathcal{O}_Y$ die Form $U = \bigcup\limits_{i\in I} U_i$ mit $U_i \in \mathcal{B}$ hat. Nach Lemma 1.3.10 ist $f^{-1}(U) = f^{-1}(\bigcup\limits_{i\in I} U_i) = \bigcup\limits_{i\in I} f^{-1}(U_i)$, also $f^{-1}(U) \in \mathcal{O}_X$ als Vereinigung offener Mengen. □

**Bemerkung 13.1.19.** Anders als bei $\mathbb{K}$-linearen Abbildungen – dort haben wir bewiesen, daß die Umkehrabbildung einer $\mathbb{K}$-linearen bijektiven Abbildung wieder $\mathbb{K}$-linear ist – muß die Umkehrabbildung einer bijektiven stetigen Abbildung nicht immer stetig sein. Das einfachste Beispiel ist: Es sei $X \neq \emptyset$ eine Menge mit mindestens zwei Elementen, dann ist $\mathrm{id}_X : (X, \mathcal{P}(X)) \to (X, \{\emptyset, X\})$ bijektiv und stetig, aber die Umkehrabbildung $\mathrm{id}_X : (X, \{\emptyset, X\}) \to (X, \mathcal{P}(X))$ ist nicht stetig, denn für $x \in X$ ist $\{x\} \in \mathcal{P}(X)$, aber $\mathrm{id}_X^{-1}(\{x\}) = \{x\} \notin \{\emptyset, X\}$.

**Lemma 13.1.20.** *Es sei $I$ eine Menge und $(X_i, \mathcal{O}_{X_i})$ seien topologische Räume für alle $i \in I$. Auf dem cartesischen Produkt $X := \prod\limits_{i\in I} X_i$ mit den kanonischen Projektionen*

$$\begin{array}{rcl} \mathrm{pr}_j : X & \to & X_j \\ (x_i)_{i\in I} & \mapsto & x_j \end{array} \quad \textit{bildet}$$

$$\mathcal{B} = \left\{ \prod_{i\in I} V_i : V_i \in \mathcal{O}_{X_i};\ \textit{es gibt endliches } J \subseteq I \textit{ mit } V_j = X_j \textit{ für alle } j \in I \backslash J \right\} \tag{13.1.11}$$

*eine Basis der* **Produkttopologie** $\mathcal{O}_X$. *$\mathcal{O}_X$ hat folgende Eigenschaften:*

*a) $\mathcal{O}_X$ ist die gröbste Topologie auf $X$, so daß alle Projektionen $\mathrm{pr}_j$, $j \in I$ stetig sind.*

*b) Ist $(Z, \mathcal{O}_Z)$ ein topologischer Raum und $g : Z \to X$ eine Abbildung, dann ist $g$ in einem Punkt $z \in Z$ genau dann stetig, wenn alle Abbildungen*

$$\mathrm{pr}_j \circ g : (Z, \mathcal{O}_Z) \to (X_j, \mathcal{O}_{X_j}), \quad j \in I$$

*in $z$ stetig sind.*

*c) Sind $(Y_i, \mathcal{O}_{Y_i})$, $i \in I$ topologische Räume und $g_i : X_i \to Y_i$ Abbildungen. Dann ist*
$$\begin{array}{rcl} \prod\limits_{i\in I} g_i : \prod\limits_{i\in I} X_i & \to & \prod\limits_{i\in I} Y_i \\ (x_i)_{i\in I} & \mapsto & (g_i(x_i))_{i\in I} \end{array}$$
*in $a = (a_i)_{i\in I} \in \prod\limits_{i\in I} X_i$ genau dann stetig (bezüglich den beiden Produkttopologien), wenn jedes $g_i : X_i \to Y_i$, $i \in I$ in $a_i \in X_i$ stetig ist.*

*Beweis.* $\mathcal{B} = \left\{ \prod\limits_{i\in I} V_i : V_i \in \mathcal{O}_{X_i};\ \text{es gibt endliches } J \subseteq I \text{ mit } V_j = X_j \text{ für alle } j \in I \backslash J \right\}$ besteht aus allen endlichen Durchschnitten von Mengen der Form $\mathrm{pr}_j^{-1}(U_j)$, $j \in I$, $U_j \in \mathcal{O}_{X_j}$. Setze

$$\mathcal{O}_X := \left\{ U \subseteq X : \text{ Es gibt eine Menge } L \text{ und } U_l \in \mathcal{B}, l \in L \text{ mit } U = \bigcup_{l\in L} U_l \right\} \cup \{\emptyset\}.$$

Wegen $X \in \mathcal{B}$ sind auch $\emptyset, X \in \mathcal{O}_X$. Offenbar sind beliebige Vereinigungen von Elementen aus $\mathcal{O}_X$ wieder in $\mathcal{O}_X$. Für eine endliche Menge $M$, oE. $M = \{1, ..., n\}$ und $U_m \in \mathcal{O}_X$, $m \in M$ gibt es $U_m = \bigcup_{l \in L_m} U_{m,l}$ mit $U_{m,l} \in \mathcal{B}$, also ergibt sich als Verallgemeinerung von (1.1.4) auf endliche Durchschitte von Vereinigungen:

$$\bigcap_{m=1}^{n} U_m = \bigcap_{m=1}^{n} \left( \bigcup_{l \in L_m} U_{m,l} \right) = \bigcup_{(l_1, ..., l_m) \in \prod_{j=1}^{n} L_j} (U_{1,l_1} \cap ... \cap U_{n,l_n}) \in \mathcal{O}_X,$$

denn $U_{1,l_1} \cap ... \cap U_{n,l_n} \in \mathcal{B}$ als endlicher Durchschnitt von Elementen aus $\mathcal{B}$. Somit ist $\mathcal{O}_X$ eine Topologie auf $X$, dazu ist $\mathcal{B}$ nach Konstruktion eine Basis der Topologie für $\mathcal{O}_X$. Da alle Mengen $\mathrm{pr}_j^{-1}(U_j), j \in I, U_j \in \mathcal{O}_{x_j}$ in $\mathcal{B}$ und damit auch in $\mathcal{O}_X$ enthalten sind, ist jede Abbildung $\mathrm{pr}_j : (X, \mathcal{O}_X) \to (X_j, \mathcal{O}_{X_j})$ stetig.
Damit jede Abbildung $\mathrm{pr}_j : (X, \mathcal{O}) \to (X_j, \mathcal{O}_{X_j})$ stetig ist, muß die Topologie $\mathcal{O}$ alle Mengen der Form $\mathrm{pr}_j^{-1}(U_j)$ mit $U_j \in \mathcal{O}_{X_j}$ enthalten. Wegen der Axiome einer Topologie, müssen auch alle endlichen Durchschnitte, also alle Elemente aus $\mathcal{B}$ in $\mathcal{O}$ enthalten sein und dann beliebige Vereinigungen von Elementen aus $\mathcal{B}$, also $\mathcal{O}_X \subseteq \mathcal{O}$. Da $\mathcal{O}_X$ bereits eine Topologie ist in der alle Projektionen $\mathrm{pr}_j, j \in J$ stetig sind, ist $\mathcal{O}_X$ die gröbste Topologie, so daß alle Abbildungen $\mathrm{pr}_j$ stetig sind.
Es sei nun $(Z, \mathcal{O}_Z)$ ein topologischer Raum und $g : Z \to X$.

„$\Rightarrow$“ Ist $g : (Z, \mathcal{O}_Z) \to (X, \mathcal{O}_X)$ stetig (in $z$), dann sind $\mathrm{pr}_j \circ g : (Z, \mathcal{O}_Z) \to (X_j, \mathcal{O}_{X_j})$ als Komposition (in $z$) stetiger Abbildungen stetig (in $z$).

„$\Leftarrow$“ Sind alle Abbildungen $\mathrm{pr}_j \circ g : Z \to X_j$ stetig in $z$ und ist $U$ eine Umgebung von $g(z)$, so gibt es ein $V \in \mathcal{O}_X$ mit $g(z) \in V \subseteq U$. Da $V$ nach Konstruktion von $\mathcal{O}_X$ eine Vereinigung von Elementen aus $\mathcal{B}$ ist, können wir oE. $V \in \mathcal{B}$ voraussetzen, dh. $V = \prod_{i \in I} V_i$ mit $V_i \in \mathcal{O}_{X_i}$ für alle $i \in I$ und es gibt eine endliche Menge $J \subseteq I$ mit $V_i = X_i$ für alle $i \in I \backslash J$. Dann ist

$$g^{-1}(V) = \left\{ w \in Z : g(w) = (g(w)_i)_{i \in I} \in V = \prod_{i \in I} V_i \right\} = \bigcap_{j \in J} (\mathrm{pr}_j \circ g)^{-1}(V_j).$$

Da $\mathrm{pr}_j \circ g$ stetig ist, folgt $(\mathrm{pr}_j \circ g)^{-1}(V_j) \in \mathcal{O}_Z$ und daher ist $g^{-1}(V) \in \mathcal{O}_Z$ als Durchschnitt von endlich vielen Elementen aus $\mathcal{O}_Z$ und da $g^{-1}(U)$ die offene Umgebung $g^{-1}(V)$ enthält, ist $g^{-1}(U)$ eine Umgebung von $z$, also $g$ stetig in $z$.

Sind nun noch $(Y_i, \mathcal{O}_{Y_i})$, $i \in I$ topologische Räume und $g_i ; X_i \to Y_i$ Abbildungen, $a = (a_i)_{i \in I} \in X = \prod_{i \in I} X_i$.

„$\Leftarrow$“ Ist $g_j : X_j \to Y_j$ stetig in $a_j$, so ist $g_j \circ \mathrm{pr}_{X,j} : X \to Y_j$ in $a = (a_i)_{i \in I}$ als Komposition der stetigen Abbildung $\begin{array}{rcl} \mathrm{pr}_{X,j} : X & \to & X_j \\ x = (x_i)_{i \in I} & \mapsto & x_j \end{array}$ mit der in $a_j$ stetigen Abbildung

$g_j : X_j \to Y_j$ wieder stetig.

$$\begin{aligned} g : X &\to \prod_{i \in I} Y_i \\ x = (x_i)_{i \in I} &\mapsto g(x) = (g_i(x_i))_{i \in I} \end{aligned}$$

ist nach (b) stetig in $a$.

„$\Rightarrow$" Die Funktion $h_j : X_j \to X$ mit $\mathrm{pr}_{X,j}(h_j(x_j)) = x_j$ und $\mathrm{pr}_{X.i}(h_j(x_j)) = a_i$ für $i \in I$, $i \neq j$ ist laut (b) stetig, da $\mathrm{id}_{X_j}$ und die konstanten Abbildungen $\begin{array}{ccc} X_j & \to & X_i \\ x_j & \to & a_i \end{array}$ stetig sind. Ist nun $g = \prod_{i \in I} g_i$ stetig in $a$, dann ist $g_i = \mathrm{pr}_{X,i} \circ g \circ h_i$ steig in $a_i$.

□

**Definition 13.1.21.** *Sind $(X_i, \mathcal{O}_i)$ topologische Räume für $i \in I$ und $X := \prod_{i \in I} X_i$, dann heißt die Topologie $\mathcal{O}_X$ aus Lemma 13.1.20 die* **Produkttopologie** *und $(X, \mathcal{O}_X)$ der* **Produktraum** *(der topologischen Räume $X_i, i \in I$).*

**Lemma 13.1.22.** *Es sei $(X, \mathcal{O}_X)$ ein topologischer Raum, $(Y, \mathcal{O}_Y)$ ein Hausdorffraum und $f : X \to Y$ und $g : X \to Y$ seien stetig. Dann gilt:*

*a) $\{x \in X : f(x) = g(x)\}$ ist eine abgeschlossene Teilmenge von $X$.*

*b) Ist $A \subseteq X$ dicht und gilt $f|_A = g|_A$, so gilt $f = g$.*

*Beweis.*

a) Es sei $\triangle(Y) := \{(y, y) \in Y \times Y\}$ und $(y_1, y_2) \in Y \times Y \setminus \triangle(Y)$, also $y_1 \neq y_2$. Da $Y$ ein Hausdorffraum ist, gibt es eine Umgebung $V_1 \in \mathcal{O}_Y$ von $y_1$ und eine Umgebung $V_2 \in \mathcal{O}_Y$ von $y_2$, so daß $V_1 \cap V_2 = \emptyset$ ist. Daher ist $V_1 \times V_2$ eine Umgebung von $(y_1, y_2)$ in $Y \times Y$ mit $(V_1 \times V_2) \cap \triangle(Y) = \emptyset$, also ist $(Y \times Y) \setminus \triangle Y$ offen oder $\triangle(Y)$ abgeschlossen. Da nach Lemma 13.1.20 die Abbildung $\begin{array}{rcl} (f, g) : X & \to & Y \times Y \\ x & \mapsto & (f(x), g(x)) \end{array}$ stetig ist, so ist

$$\{x \in X : f(x) = g(x)\} = (f, g)^{-1}(\triangle(Y))$$

abgeschlossen.

b) Nach Voraussetzung ist $\{x \in X : f(x) = g(x)\} \supseteq A$ und $\overline{A} = X$. Da nach Teil a) die Menge $\{x \in X : f(x) = g(x)\}$ abgeschlossen ist, folgt $\{x \in X : f(x) = g(x)\} \supseteq \overline{A} = X$.

□

**Korollar 13.1.23.** *Es seien $(X, \mathcal{O}_X)$ ein topologischer Raum und $(Y, \mathcal{O}_Y)$ ein Hausdorffraum, $f : X \to Y$ sei stetig, dann ist der Graph von $f$*

$$\Gamma(f) := \{(x, f(x)) : x \in X\} \tag{13.1.12}$$

*eine abgeschlossene Teilmenge von $X \times Y$.*

*Beweis.* Nach Definition der Produkttopologie sind die Projektionen $\begin{array}{rcl}\mathrm{pr}_1 : X \times Y & \to & X \\ (x,y) & \mapsto & x\end{array}$ und $\begin{array}{rcl}\mathrm{pr}_2 : X \times Y & \to & Y \\ (x,y) & \mapsto & y\end{array}$ stetig. Deshalb ist nach Satz 13.1.17 auch $f \circ \mathrm{pr}_1$ stetig und daher

$$\Gamma(f) = \{(x, f(x)) : x \in X\} = \{(x,y) \in X \times Y : \mathrm{pr}_2((x,y)) = (f \circ \mathrm{pr}_1)(x,y)\}$$

in $X \times Y$ abgeschlossen nach Satz 13.1.15. □

**Definition 13.1.24.** *Es seien $(X, \mathcal{O}_X)$ und $(Y, \mathcal{O}_Y)$ topologische Räume, $A \subseteq X$, $a \in X$ ein Berührpunkt [2] von $A$ und $f : A \to Y$ eine Abbildung. Falls*

- *$a \notin A$, dann heißt $y \in Y$ der* **Grenzwert von $f$ für gegen $a$ strebendes $x \in A$**, *in Zeichen*

  $$y = \lim_{\substack{x \to a \\ x \in A}} f(x)$$

  *wenn die Funktion* $\begin{array}{rcl} g : A \cup \{a\} & \to & Y \\ x & \mapsto & \begin{cases} f(x) & \text{für } x \in A \\ y & \text{für } x = a \end{cases}\end{array}$ *im Punkt $a$ stetig ist.*

- *$a \in A$, dann heißt $y \in Y$ der* **Grenzwert von $f$ für gegen $a$ strebendes $x \in A$**, *in Zeichen*

  $$y = \lim_{\substack{x \to a \\ x \in A}} f(x)$$

  *genau dann, wenn $y = f(a)$ und $f$ in $a$ stetig ist.*

## 13.2 Metrische Räume: Stetige und gleichmäßig stetige Abbildungen

**Satz 13.2.1.** *Es seien $(X, d)$ und $(Y, d')$ metrische Räume, $f : X \to Y$ und $a \in X$. Dann sind äquivalent:*

*a) Für jedes $\varepsilon > 0$ gibt es ein $\delta = \delta(a, \varepsilon) > 0$ mit*

$$d'(f(x), f(a)) < \varepsilon \quad \text{für alle} \quad x \in X \text{ mit } d(x,a) < \delta. \tag{13.2.1}$$

*b) Für jede Folge $(x_n)_{n \in \mathbb{N}}$ in $X$ mit $\lim_{n \to \infty} x_n = a$ konvergiert auch die Folge $(f(x_n))_{n \in \mathbb{N}}$ in $Y$ und es gilt:*

$$\lim_{n \to \infty} f(x_n) = f(a) \tag{13.2.2}$$

[2] ist $a$ kein Berührpunkt von $A$, so gibt es eine Umgebung $U$ von $a$ mit $U \cap A = \emptyset$ – dann kann man nicht mit $x \in A$ gegen $a$ gehen

*c) $f : X \to Y$ ist als Abbildung zwischen den topologischen Räumen $(X, \mathcal{O}_d)$ und $(Y, \mathcal{O}_{d'})$ stetig in $a$.*

*Beweis.*

a)⇒ b) Es sei $a \in X$ und $\varepsilon > 0$, dann gibt es ein $\delta = \delta(a, \varepsilon) > 0$ mit $d'(f(a), f(x)) < \varepsilon$ für alle $x \in X$ mit $d(a, x) < \delta$. Daher gibt es für jede Folge $(x_n)_{n\in\mathbb{N}}$ in $X$ mit $\lim_{n\to\infty} x_n = a$ ein $N = N(\varepsilon) \in \mathbb{N}$ mit $d(x_n, a) < \delta(a, \varepsilon)$ für $n \geq N(\varepsilon)$, also ist $d'(f(x_n), f(a)) < \varepsilon$ für alle $n \geq N(\varepsilon)$. Da $\varepsilon > 0$ beliebig war, konvergiert $(f(x_n))_{n\in\mathbb{N}}$ in $(Y, d')$ mit $f(a) = \lim_{n\to\infty} f(x_n)$.

b)⇒c) (durch $\neg$ c) ⇒ $\neg$ b)) Betrachte die Umgebungsbasis $\{K(a, \frac{1}{n}) : n \in \mathbb{N}\}$ von $a$ in $\mathcal{O}_d$. Angenommen $f$ ist in $a$ nicht stetig, dann existiert eine Umgebung $V$ von $f(a)$, so daß $f^{-1}(V)$ keine Umgebung von $a$ ist; da $\{K(a, \frac{1}{n}) : n \in \mathbb{N}\}$ Umgebungsbasis von $a$ ist, folgt $f^{-1}(V) \not\supseteq K(a, \frac{1}{n})$ für alle $n \in \mathbb{N}$. Deshalb existiert für jedes $n \in \mathbb{N}$ ein $x_n \in K(a, \frac{1}{n})$ mit $f(x_n) \notin V$. Deshalb gilt $\lim_{n\to\infty} x_n = a$ und $(f(x_n))_{n\in\mathbb{N}}$ konvergiert nicht gegen $f(a)$.

c)⇒a) Es sei $\varepsilon > 0$, dann ist $K(f(a), \varepsilon) \in \mathcal{O}_{d'}$ eine Umgebung von $f(a)$, also $f^{-1}(K(f(a), \varepsilon))$ eine Umgebung von $a$, also gibt es nach Definition von $\mathcal{O}_d$ ein $\delta = \delta(a, \varepsilon) > 0$ mit $K(a, \delta(a, \varepsilon)) \subseteq f^{-1}(K(f(a), \varepsilon)) = \{x \in X : d'(f(x), f(a)) < \varepsilon\}$. Daher ist $d'(f(a), f(x)) < \varepsilon$ für alle $x \in X$ mit $d(x, a) < \delta(a, \varepsilon)$. □

**Korollar 13.2.2.** *Sind $(X, d)$ und $(Y, d')$ metrische Räume, $A \subseteq X$, $a \in \overline{A}$ und $f : A \to Y$ eine Abbildung. Dann sind äquivalent:*

*a) $\lim_{\substack{x\to a \\ x\in A}} f(x)$ existiert und $y = \lim_{\substack{x\to a \\ x\in A}} f(x)$*

*b) Für jede Folge $(x_n)_{n\in\mathbb{N}}$ in $X$ mit $x_n \in A$ und $\lim_{n\to\infty} x_n = a$ konvergiert $(f(x_n))_{n\in\mathbb{N}}$ in $Y$ und $y = \lim_{n\to\infty} f(x_n)$.*

*Beweis.* Das Korollar formuliert die Aussage von Satz 13.2.1 durch Grenzwerte. □

**Definition 13.2.3.** *Sind $(X, d)$ und $(Y, d')$ metrische Räume, dann heißt eine Abbildung $f : X \to Y$*

- **gleichmäßig stetig**, *wenn es für alle $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ gibt mit*

$$d'(f(x), f(y)) < \varepsilon \quad \text{für alle } x, y \in X \text{ mit } d(x, y) < \delta. \tag{13.2.3}$$

- *eine* **Kontraktion**, *wenn es ein $q \in ]0, 1[$ gibt, so daß*

$$d'(f(x), f(y)) \leq q d(x, y) \quad \text{für alle } x, y \in X. \tag{13.2.4}$$

**Bemerkung 13.2.4.** Sind $(X, d)$ und $(Y, d')$ metrische Räume und ist $f : X \to Y$ eine Kontraktion, so ist $f$ auch gleichmäßig stetig und jede gleichmäßig stetige Funktion $g : X \to Y$ ist auch stetig.

**Satz 13.2.5.** *Es sei $(V, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum. $\mathbb{K}$ und $V$ seien mit den Normtopologien versehen und $\mathbb{K}\times V$ und $V\times V$ mit den Produkttopologien. Dann*

- *sind die Additionen $+ : V\times V \to V$ und $+ : \mathbb{K}\times\mathbb{K}\to\mathbb{K}$ gleichmäßig stetig zB. mit der Norm* $\begin{array}{rcl} |||\cdot||| : V\times V & \to & [0,\infty[ \\ (v,w) & \mapsto & \max\{\|v\|,\|w\|\} \end{array}$ *auf $V\times V$.*
- *ist die Skalarmultiplikation $\cdot : \mathbb{K}\times V \to V$ stetig*
- *ist die Multiplikation $\cdot : \mathbb{K}\times\mathbb{K}\to\mathbb{K}$ stetig*
- *ist die Norm $\|\cdot\| : V \to [0,\infty[$ gleichmäßig stetig.*

*Beweis.*

- Es sei $|||\cdot|||$ obige Norm auf dem cartesischen Produkt [3] . Sind $(x_1,x_2),(y_1,y_2)\in V\times V$, dann folgt aus der Dreiecksungleichung für $\|\cdot\|$ die Abschätzung

$$\begin{aligned} &\|[+(x_1,x_2)]-[+(y_1,y_2)]\| = \|x_1+x_2-y_1-y_2\| = \|(x_1-y_1)+(x_2-y_2)\| \\ &\quad\leq\ \|x_1-y_1\|+\|x_2-y_2\| \leq 2|||(x_1,x_2)-(y_1,y_2)|||, \end{aligned}$$

  also ist $+ : V\times V\to V$ gleichmäßig stetig und in der Definition läßt sich $\delta(\varepsilon)=\frac{\varepsilon}{2}$ wählen.

- Nach Korollar 5.4.10 und Satz 13.2.1 ist die Skalarmultiplikation stetig.

- Aus der Dreiecksungleichung folgt für jedes $x,y\in V$:

$$\begin{aligned} \|x\| &= \|x-y+y\| \leq \|x-y\|+\|y\| \\ \|y\| &= \|y-x+x\| \leq \|y-x\|+\|x\| \end{aligned}$$

  woraus

$$\begin{aligned} \|x\|-\|y\| &\leq \|x-y\| \\ \|y\|-\|x\| &\leq \|y-x\| \end{aligned}$$

  oder

$$\big|\|x\|-\|y\|\big| \leq \|x-y\| \tag{13.2.5}$$

  folgt. Wegen (13.2.5) ist die Norm $\|\cdot\| : V\to[0,\infty[$ gleichmäßig stetig, wobei man $\delta(\varepsilon)=\varepsilon$ in der Definition wählen kann.

Die beiden anderen Behauptungen ergeben sich für den Spezialfall $V=\mathbb{K}$. □

**Korollar 13.2.6.** *Ist $V$ ein normierter $\mathbb{K}-$ Vektorraum, dann ist zu $a_n,\dots,a_1,a_0\in V$ die Polynomfunktion* $\begin{array}{rcl} f:\mathbb{K} & \to & V \\ x & \mapsto & a_nx^n+\dots+a_1x+a_0 \end{array}$ *(bzgl. der Standardtopologie $\mathcal{O}_{\mathbb{K}}$ auf $\mathbb{K}$ und der Normtopologie auf $V$) stetig.*

[3] dann ist die Normtopologie $\mathcal{O}_{|||\cdot|||}$, die von $|||\cdot|||$ erzeugt wird dasselbe wie die Produkttopologie von $\mathcal{O}_{\|\cdot\|}$ mit $\mathcal{O}_{\|\cdot\|}$

*Beweis.* $\mathrm{id}_{\mathbb{K}}$ ist offenbar gleichmäßig stetig, daher ist $\begin{array}{rcl}\tau_2 : \mathbb{K} & \to & \mathbb{K} \\ x & \mapsto & x^2\end{array} = \cdot \circ (\mathrm{id}_{\mathbb{K}}, \mathrm{id}_{\mathbb{K}})$ nach Lemma 13.1.20 über die Produkttopologie und Satz 13.2.5 als Komposition stetiger Abbildungen stetig. Mit Induktion zeigt man dann, daß alle Monome $\begin{array}{rcl}\tau_n : \mathbb{K} & \to & \mathbb{K} \\ x & \mapsto & x^n\end{array}$ wegen $\tau_n = \cdot \circ (\mathrm{id}_{\mathbb{K}}, \tau_{n-1})$ stetig sind. Da auch die Skalarmultiplikation stetig ist, ist $a_k\tau_k$ stetig und damit ist jede Polynomfunktion stetig als Summe von Monomen. □

**Korollar 13.2.7.** *Ist $(V, \langle\cdot,\cdot\rangle)$ ein $\mathbb{K}$-Vektorraum mit Skalarprodukt, dann ist bezüglich den Normtopologien auf $V$ bzw. $\mathbb{K}$ das Skalarprodukt $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}$ stetig.*

*Beweis.* Folgt durch Anwenden der Polarisierungsidentität aus der Stetigkeit der Norm. □

**Lemma 13.2.8.** *Es sei $(X, d)$ ein metrischer Raum und $\emptyset \neq Y \subseteq X$. Dann ist*

$$\begin{array}{rcl}\mathrm{dist}(\cdot, Y) : X & \to & [0, \infty[ \\ x & \mapsto & \mathrm{dist}(x, Y) := \inf\{d(x, y) : y \in Y\}\end{array}$$

*gleichmäßig stetig.*

*Beweis.* Tutorium □

**Satz 13.2.9.** *Sei $(X, d)$ ein metrischer Raum, $x \in X$ und $\emptyset \neq A \subseteq X$. $\overline{A}$ sei der Abschluß von $A$ und* $\mathrm{dist}(x, A) := \inf\{d(x, y) : y \in A\}$.[4] *Dann sind äquivalent:*

*a)* $x \in \overline{A}$.

*b) Es gibt eine Folge $(x_n)_{n\in\mathbb{N}}$ in $A$ mit $x = \lim_{n\to\infty} x_n$.*

*c)* $\mathrm{dist}(x, A) = 0$.

*Beweis.*

a) ⇒ b) Da $(X, d)$ ein metrischer Raum ist, ist für jedes $x \in X$ das System

$$\mathcal{U}(x) := \{K(x, \frac{1}{n}) : n \in \mathbb{N}\}$$

eine abzählbare Umgebungsbasis von $x$. Ist also $x$ ein Berührpunkt von $A$, so gilt $K(x, \frac{1}{n}) \cap A \neq \emptyset$ und daher gibt es für jedes $n \in \mathbb{N}$ ein $x_n \in K(x, \frac{1}{n}) \cap A$ und b) folgt.

b) ⇒ c) Ist $(x_n)_{n\in\mathbb{N}}$ eine Folge in $A$ mit $x = \lim_{n\to\infty} x_n$, dann gibt es für jedes $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$ mit $x_n \in K(x, \varepsilon)$ für $n \geq N(\varepsilon)$. Da $x_n \in A$ ist, folgt

$$0 \leq \mathrm{dist}(x, A) \leq \sup\{d(x, x_n) : n \geq N(\varepsilon)\} \leq \varepsilon$$

und nach Bilden von $\inf_{\varepsilon>0}$ folgt $0 \leq \mathrm{dist}(x, A) \leq 0$.

[4] dann ist $\emptyset \neq \{d(x, y) : y \in A\} \subseteq \mathbb{R}$ zB. durch 0 nach unten beschränkt, also existiert $\inf\{d(x, y) : y \in A\}$

c) $\Rightarrow$ a) Ist $\operatorname{dist}(x, A) = 0$, so gibt es wegen der inf-Definition eine Folge $(x_n)_{n\in\mathbb{N}}$ in $A$ mit $\inf\{d(x,y) : y \in A\} = \lim\limits_{n\to\infty} d(x, x_n)$. Ist $U$ eine Umgebung von $x$, so gibt es $\varepsilon > 0$ mit $K(x,\varepsilon) \subseteq U$ und somit gibt es wegen $\lim\limits_{n\to\infty} d(x, x_n) = 0$ ein $N(\varepsilon) \in \mathbb{N}$ mit $x_n \in K(x,\varepsilon)$ für alle $n \geq N(\varepsilon)$, daher ist $x \in \overline{A}$. □

**Korollar 13.2.10.** *Ist $(X, d)$ ein vollständiger metrischer Raum, $A \subseteq X$ abgeschlossen, so ist $A$ vollständig.*

*Beweis.* Ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $A$, so existiert $x = \lim\limits_{n\to\infty} x_n \in X$, da $X$ vollständig ist. Da $A$ abgeschlossen liegt auch der Grenzwert $x$ der Folge $(x_n)_{n\in\mathbb{N}}$ nach Satz 13.2.9 in $A$, also ist $A$ vollständig. □

## 13.3 Banachscher Fixpunktsatz

**Satz 13.3.1 (Fixpunktsatz von Banach).** *Sei $(X, d)$ ein vollständiger metrischer Raum, $X \neq \emptyset$ und $T : X \to X$ eine Abbildung, so daß ein $q \in ]0, 1[$ existiert mit*

$$d(T(x), T(y)) \leq q d(x, y) \tag{13.3.1}$$

*für alle $x, y \in X$ (**Kontraktion**), dann gilt:*

1. *$T$ hat genau einen **Fixpunkt** $a \in X$, d.h. es gibt genau ein $a \in X$ mit $a = T(a)$.*
2. *Konstruktives Verfahren: Für beliebiges $x_0 \in X$ setze $x_{n+1} := T(x_n)$ für $n \in \mathbb{N}$, dann gilt $\lim\limits_{n\to\infty} x_n = a$.*
3. *Fehlerabschätzungen: Für $n \in \mathbb{N}_0$ gilt:*

$$d(x_n, a) \leq \frac{q^n}{1-q} d(x_0, x_1) \tag{13.3.2}$$

$$d(x_{n+1}, a) \leq \frac{q}{1-q} d(x_n, x_{n+1}) \tag{13.3.3}$$

$$d(x_{n+1}, a) \leq q d(x_n, a) \tag{13.3.4}$$

*Beweis.*
Eindeutigkeit: Sind $a, b \in X$ Fixpunkte von $T$, dann ist $d(a,b) = d(T(a), T(b)) \leq q d(a,b)$, woraus wegen $q < 1$ dann $d(a,b) = 0$, also $a = b$ folgt.
Existenz: Ist $x_0 \in X$, $x_{n+1} := T(x_n)$ für $n \geq 1$ und $k \leq m$, dann gilt:

$$\begin{aligned} d(x_k, x_m) &\leq q d(x_{k-1}, x_{m-1}) \leq \ldots \leq q^k d(x_0, x_{m-k}) \leq q^k \sum_{j=0}^{m-k-1} d(x_j, x_{j+1}) \leq \\ &\leq d(x_0, x_1) q^k \sum_{j=0}^{m-k-1} q^j = d(x_0, x_1) q^k \frac{1-q^{m-k}}{1-q} \leq d(x_0, x_1) \frac{q^k}{1-q}, \end{aligned} \tag{13.3.5}$$

also ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $(X,d)$. Weil $(X,d)$ vollständig ist, existiert der Grenzwert $a := \lim\limits_{n\to\infty} x_n \in X$. Aufgrund der Kontraktionseigenschaft (13.3.1) ist $T$ (gleichmäßig) stetig, daher

$$a = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} T(x_n) = T(\lim_{n\to\infty} x_n) = T(a),$$

also $a$ Fixpunkt von $T$. (13.3.2) folgt aus (13.3.5) mit $k = n$, $m = n+l$ für $l \to \infty$.

$$\begin{aligned} d(x_{n+1}, x_{n+m+1}) &\le d(x_{n+1}, x_{n+2}) + \dots + d(x_{n+m}, x_{n+m+1}) \le \\ &\le d(x_n, x_{n+1})(q + q^2 + \dots + q^m) \le d(x_n, x_{n+1})\frac{q}{1-q} \end{aligned}$$

ergibt (13.3.3) im Limes $m \to \infty$. $d(x_{n+1}, a) = d(T(x_n), T(a)) \le qd(x_n, a)$ ist dann (13.3.4). □

**Beispiel 13.3.2.** $\begin{array}{rcl} T : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^2 \end{array}$ ist wegen $|x^2 - y^2| = |x+y| \cdot |x-y|$ keine Kontraktion, wohl aber $\begin{array}{rcl} t : [-\frac{1}{4}, \frac{1}{4}] & \to & [-\frac{1}{4}, \frac{1}{4}] \\ x & \mapsto & x^2 \end{array}$ und [5] Korollar 13.2.10 zeigt dann, daß $[-\frac{1}{4}, \frac{1}{4}]$ abgeschlossen ist, man dann den Banachschen Fixpunktsatz anwenden kann. Auch wenn man hier sofort $x = 0$ als Fixpunkt nachrechnet, zeigt der Fixpunktsatz daß dies der einzige Fixpunkt von $t$ in $[-\frac{1}{4}, \frac{1}{4}]$ ist.

## 13.4 Lokal gleichmäßige Konvergenz von Funktionenfolgen und (Potenz-)Reihen

**Definition 13.4.1.** *Es sei $(X, \mathcal{O}_X)$ ein topologischer Raum und $(Y, d)$ ein metrischer Raum. Eine Folge von Funktionen $(f_n : X \to Y)_{n\in\mathbb{N}}$ heißt*

- **punktweise konvergent** *gegen $f : X \to Y$, wenn für jedes $x \in X$ gilt:*

$$\lim_{n\to\infty} d(f_n(x), f(x)) = 0 \tag{13.4.1}$$

- **gleichmäßig konvergent** *gegen $f : X \to Y$, wenn*

$$\lim_{n\to\infty} \Big(\sup\{d(f_n(x), f(x)) : x \in X\}\Big) = 0 \tag{13.4.2}$$

- **lokal gleichmäßig konvergent** *gegen $f : X \to Y$, wenn es für jedes $x \in X$ eine Umgebung $U(x)$ von $x$ gibt, so daß $(f_n|_{U(x)} : U(x) \to Y)_{n\in\mathbb{N}}$ gleichmäßig gegen $f|_{U(x)}$ konvergiert.*

**Bemerkung 13.4.2.** Ist $(X, \mathcal{O}_X)$ ein topologischer Raum, $(Y, d)$ ein metrischer Raum und konvergiert $f_n : X \to Y$ gleichmäßig gegen $f : X \to Y$, dann konvergiert $f_n$ auch lokal gleichmäßig und punktweise gegen $f$. Als Kandidat für den (lokal) gleichmäßigen Grenzwert kommt also immer nur der punktweise Grenzwert in Frage. Dann muß man

[5] für $|x| \le \frac{1}{4}$ ist $x^2 \le |x| \le \frac{1}{4}$, also $t$ wohldefiniert

aber noch die (lokal) gleichmäßige Konvergenz überprüfen, denn nicht jeder punktweise Grenzwert ist auch (lokal) gleichmäßiger Grenzwert, wie das folgende Beispiel zeigt:

$$\begin{array}{rcl} f_n : [0,1] & \to & \mathbb{R} \\ x & \mapsto & x^n \end{array}$$

- konvergiert punktweise gegen $\begin{array}{rcl} f : [0,1] & \to & \mathbb{R} \\ x & \mapsto & \begin{cases} 0 & \text{falls } x \in [0,1[ \\ 1 & \text{falls } x = 1 \end{cases} \end{array}$
- konvergiert nicht lokal gleichmäßig, denn jede Umgebung von 1 (in der Relativtopologie von $[0,1]$) enthält ein Intervall $]1-\delta,1]$ mit geeignetem $\delta > 0$. Da die Funktionen $f_n$ streng monoton steigend und stetig sind, ist $\max\limits_{x\in[1-\delta,1]} f_n(x) = \lim\limits_{x\to 1} f_n(x) = f_n(1) = 1$, also gilt für alle $n \in \mathbb{N}$:

  $$\sup\{|f_n(x) - f(x)| : x \in ]1-\delta,1]\} \geq \sup\{|f_n(x)| : x \in ]1-\delta,1[\} = 1$$

  und damit konvergiert $f_n$ nicht lokal gleichmäßig gegen $f$.

**Satz 13.4.3.** *Es sei $X$ ein topologischer Raum und $(Y,d)$ ein metrischer Raum, $f_n : X \to Y$ seien stetig und lokal gleichmäßig konvergent mit Grenzwert $f : X \to Y$, dann ist $f$ stetig.*

*Beweis.* Es sei $a \in X$ und $U$ eine Umgebung von $a$, so daß $f_n|_U$ gleichmäßig gegen $f|_U$ konvergiert, dh. für jedes $\varepsilon > 0$ existiert ein $N = N(\varepsilon) \in \mathbb{N}$ mit $\sup\limits_{x\in U} d(f_n(x), f(x)) < \dfrac{\varepsilon}{3}$ für alle $n \geq N(\varepsilon)$. Wegen Stetigkeit von $f_N$ im Punkt $a$ gibt es eine Umgebung $V \subseteq U$ von $a$ mit $d(f_N(a), f_N(x)) < \frac{\varepsilon}{3}$ für alle $x \in V$. Damit folgt für alle $x \in V$:

$$d(f(a), f(x)) \leq d(f(a), f_N(a)) + d(f_N(a), f_N(x)) + d(f_N(x), f(x)) \leq \varepsilon$$

und somit ist $f$ stetig im Punkt $a$. $\square$

**Satz 13.4.4.** *Ist $(X, \|\cdot\|)$ ein Banachraum und $\rho > 0$ der Konvergenzradius von $\left(\sum\limits_{n=0}^{N} a_n(z-a)^n\right)_{N\in\mathbb{N}_0}$, dann konvergiert die Folge der Partialsummen*

$$\begin{array}{rcl} f_N : \{z \in \mathbb{C} : |z-a| < \rho\} & \to & X \\ z & \mapsto & \sum\limits_{n=0}^{N} a_n(z-a)^n \end{array} \tag{13.4.3}$$

*lokal gleichmäßig, daher ist die Grenzfunktion*

$$\begin{array}{rcl} f : \{z \in \mathbb{C} : |z-a| < \rho\} & \to & X \\ z & \mapsto & \sum\limits_{n=0}^{\infty} a_n(z-a)^n \end{array} \tag{13.4.4}$$

*stetig.*

*Beweis.* $\left(\sum_{n=0}^{N} \|a_n\|(z-a)^n\right)_{N\in\mathbb{N}_0}$ hat als komplexe Potenzreihe den Konvergenzradius

$$\rho := \frac{1}{\overline{\lim\limits_{n\to\infty}} \sqrt[n]{\|a_n\|}},$$

daher konvergiert $\sum_{n=0}^{\infty} \|a_n\||z-a|^n$ für alle $z \in \mathbb{C}$ mit $|z-a| < \rho$ und somit ist $\sum_{n=0}^{\infty} a_n(z-a)^n$ absolut konvergent, und da $X$ vollständig ist auch konvergent. Zu $z \in \mathbb{C}$ mit $|z-a| < \rho$ wähle $r := \frac{1}{2}(\rho - |z-a|) > 0$ , dann ist

$$|w-a| \leq |w-z| + |z-a| \leq r + |z-a| = \frac{\rho}{2} + \frac{|z-a|}{2} =: \sigma_z < \rho$$

für $w \in \mathbb{C}$ mit $|z-w| < r$, dh. $\{w \in \mathbb{C} : |w-z| < r\} \subseteq \{u \in \mathbb{C} : |u-a| < \rho\}$. Die Partialsummenfolge $\begin{array}{rcl} f_N : \{z \in \mathbb{C} : |z-a| < \rho\} & \to & X \\ z & \mapsto & \sum_{n=0}^{N} a_n(z-a)^n \end{array}$ besteht aus stetigen Funktionen, die nach dem eben bewiesenen punktweise gegen $f$ konvergieren. Für $w \in \{w \in \mathbb{C} : |w-z| < r\}$ ist

$$\begin{aligned} \|f_N(w) - f(w)\| &= \left\|\sum_{n=N+1}^{\infty} a_n(w-a)^n\right\| \leq \sum_{n=N+1}^{\infty} \|a_n\||w-a|^n \leq \sum_{n=N+1}^{\infty} \|a_n\|(\sigma_z)^n \\ &= \sum_{n=N+1}^{\infty} (\|a_n\|\tau^n)\left(\frac{\sigma_z}{\tau}\right)^n \end{aligned}$$

mit $\sigma_z < \tau < \rho$. Nach Definition von $\rho$ erhalten wir (unabhängig von $w$) ein $C < \infty$ mit $\|a_n\|\tau^n \leq C$ für alle $n \in \mathbb{N}$, damit [6] ist

$$\sup_{w\in\mathbb{C},|w-z|<r} \|f_N(w) - f(w)\| \leq C \sum_{n=N+1}^{\infty} \left(\frac{\sigma_z}{\tau}\right)^n = C\left(\frac{\sigma_z}{\tau}\right)^{N+1} \frac{1}{1-\frac{\sigma_z}{\tau}}$$

dh. $f_N|_{K(z,r)}$ konvergiert gleichmäßig gegen $f|_{K(z,r)}$ und damit folgt die Stetigkeit von $f$ aus Satz 13.4.3. □

**Satz 13.4.5** (Weierstraßsches Majorantenkriterium)**.** *Sei $(X, \mathcal{O}_X)$ ein topologischer Raum, $(Y, \|\cdot\|)$ ein Banachraum, und $(f_i : X \to Y)_{i\in I}$ eine Funktionenfamilie und $c_i \in [0,\infty[$ mit*

*a)* $\sup\{\|f_i(x)\| : x \in X\} \leq c_i$

*b) $(c_i)_{i\in I}$ ist eine absolut summierbare Familie.*

*Dann ist die Funktionenfamilie $(f_i)_{i\in I}$ absolut und gleichmäßig summierbar.*

[6] denn nach Charakterisierung des $\limsup$ gibt es für alle $\varepsilon > 0$ ein $M(\varepsilon)$ mit $\sqrt[m]{\|a_m\|} < \limsup_{n\to\infty} \sqrt[n]{\|a_n\|} + \varepsilon = \frac{1}{\rho} + \varepsilon$ für alle $m \geq M(\varepsilon)$. Wegen $\tau < \rho$ läßt sich $\varepsilon \in ]0, \frac{1}{\tau} - \frac{1}{\rho}[$ wählen und dann ist $\sqrt[m]{\|a_m\|}\tau < \frac{\tau}{\rho} + \varepsilon\tau < \tau(\frac{1}{\rho} + \frac{1}{\tau} - \frac{1}{\rho}) = 1$, also $\|a_m\|\tau^m \leq 1$ für $m \geq M(\varepsilon)$

*Beweis.* Da $(c_i)_{i\in I}$ (absolut) summierbar ist, ist $\left(\sum_{i\in H} c_i\right)_{H\in\mathcal{E}(I)}$ ein Cauchynetz, also

$$\sup\left\{\sum_{i\in H}|c_i| : H\in\mathcal{E}(I)\right\} = \sup\left\{\sum_{i\in H}c_i : H\in\mathcal{E}(I)\right\} < \infty.$$

Für jedes $x\in X$ und $i\in I$ ist $\|f_i(x)\|\leq c_i$, also

$$\sup\left\{\sum_{i\in H}\|f_i(x)\| : H\in\mathcal{E}(I)\right\} \leq \sup\left\{\sum_{i\in H}c_i : H\in\mathcal{E}(I)\right\} < \infty$$

und damit $(f_i(x))_{i\in I}$ absolut summierbare Familie in $Y$. Das Partialsummennetz

$$\left(\begin{array}{rcl}\sum_{i\in H} f_i : X & \to & Y \\ x & \mapsto & \sum_{i\in H} f_i(x)\end{array}\right)_{H\in\mathcal{E}(I)}$$

konvergiert punktweise absolut und laut Lemma 6.4.15 auch punktweise mit dem Grenzwert

$$\begin{array}{rcl} f : X & \to & Y \\ x & \mapsto & f(x) = \sum_{i\in I} f_i(x)\end{array}$$

Da $(c_i)_{i\in I}$ (absolut) summierbar ist, gibt es für jedes $\varepsilon>0$ ein $H(\varepsilon)\in\mathcal{E}(I)$ mit $\left\|\sum_{i\in L}c_i\right\| < \varepsilon$ für alle $L\in\mathcal{E}(I)$ mit $L\cap H(\varepsilon)=\emptyset$. Weil $f: X\to Y$ punktweiser Grenzwert ist, existiert für alle $x\in X$ ein $H(\varepsilon,x)\in\mathcal{E}(I)$ mit

$$\left\|f(x) - \sum_{i\in K} f_i(x)\right\| < \varepsilon \tag{13.4.5}$$

für alle $K\in\mathcal{E}(I)$ mit $H(\varepsilon,x)\subseteq K$. Da (13.4.5) insbesondere für $K\in\mathcal{E}(I)$ mit $H(\varepsilon)\cup H(\varepsilon,x)\subseteq K$ erfüllt ist, dürfen wir oE. $H(\varepsilon)\subseteq H(\varepsilon,x)$ voraussetzen. Für jedes $L\in\mathcal{E}(I)$ mit $H(\varepsilon)\subseteq L$ gilt

$$\begin{aligned}\left\|f(x) - \sum_{i\in L} f_i(x)\right\| &\leq \left\|f(x) - \sum_{i\in K_x} f_i(x)\right\| + \left\|\sum_{i\in K_x} f_i(x) - \sum_{i\in L} f_i(x)\right\| \\ &\leq \left\|f(x) - \sum_{i\in K_x} f_i(x)\right\| + \sum_{i\in(K_x\cup L)\setminus(K_x\cap L)} \|f_i(x)\| \leq 2\varepsilon, \end{aligned} \tag{13.4.6}$$

für jedes $K_x\in\mathcal{E}(I)$ mit $H(\varepsilon,x)\subseteq K_x$ nach (13.4.5) und da $(K_x\cup L)\setminus(K_x\cap L)\in\mathcal{E}(I)$ mit $H(\varepsilon)\cap(K_x\cup L)\setminus(K_x\cap L)=\emptyset$. Da

$$\left\|f(x) - \sum_{i\in L} f_i(x)\right\| < 2\varepsilon$$

– evtl. mit verschiedenen $K_x$ bewiesen – für alle $x \in X$ gilt, ist

$$\sup\left\{\left\|f(x) - \sum_{i\in L} f_i(x)\right\| : x \in X\right\} \leq 2\varepsilon \tag{13.4.7}$$

für alle $L \in \mathcal{E}(I)$ mit $H(\varepsilon) \subseteq L$, dh. $\left(\sum_{i\in H} f_i\right)_{H\in\mathcal{E}(I)} \to f$ konvergiert gleichmäßig. □

## 13.5 Zusammenhang und Zwischenwertsatz

**Definition 13.5.1.** *Ein topologischer Raum $(X, \mathcal{O})$ heißt* **zusammenhängend**, *wenn $X$ nicht die Vereinigung disjunkter, nichtleerer offener Mengen ist. Ist $Y \subseteq X$, dann heißt $Y$ zusammenhängend, wenn der topologische Raum $(Y, \mathcal{O}_Y)$ mit der Relativtopologie $\mathcal{O}_Y$ zusammenhängend ist.*

**Lemma 13.5.2.** *Es sei $(X, \mathcal{O})$ ein topologischer Raum, dann sind äquivalent:*

*a) $X$ ist zusammenhängend*

*b) $X$ ist nicht die Vereinigung disjunkter, nichtleerer abgeschlossener Mengen*

*c) Ist $\emptyset \neq W \subseteq X$ sowohl offen als auch abgeschlossen bzgl. $\mathcal{O}$, so gilt $W = X$.*

*Beweis.*

a)⇒b) Ist $X = A_1 \cup A_2$ mit $A_1, A_2 \neq \emptyset$, $A_1 \cap A_2 = \emptyset$ und abgeschlossenen Mengen $A_1, A_2$, dann ist auch $A_1 = X\backslash A_2$ und $A_2 = X\backslash A_1$ offen, daher ist $X$ nicht zusammenhängend.

b)⇒c) Ist $\emptyset \neq W \subseteq X$ offen und abgeschlossen, dann ist $W \cup (X\backslash W)$ eine Zerlegung von $X$ in disjunkte, abgeschlossene Mengen.

c)⇒a) Ist $X = U_1 \cup U_2$ mit offenen Mengen mit $U_1 \cap U_2 = \emptyset$, so sind auch $U_1 = X\backslash U_2$ und $U_2 = X\backslash U_1$ abgeschlossen, so folgt nach der Voraussetzung in c), daß $U_1 = X$, $U_2 = \emptyset$ oder $U_1 = \emptyset$, $U_2 = X$, dh. $X$ ist zusammenhängend. □

**Satz 13.5.3.** *Es seien $(X, \mathcal{O}_X)$ und $(Y, \mathcal{O}_Y)$ topologische Räume und $f : X \to Y$ sei stetig und $Z \subseteq X$ sei zusammenhängend. Dann ist auch $f(Z) \subseteq Y$ zusammenhängend.*

*Beweis.* Angenommen $f(Z)$ ist nicht zusammenhängend, dann gibt es offene Mengen $U_1, U_2 \in \mathcal{O}_{f(Z)}$ mit $U_1, U_2 \neq \emptyset$, $U_1 \cap U_2 = \emptyset$ und $f(Z) = U_1 \cup U_2$. Dann sind $U_i = V_i \cap f(Z)$ mit $V_i \in \mathcal{O}_Y$ und

- $f^{-1}(V_1), f^{-1}(V_2) \in \mathcal{O}_X$ als Urbilder der offenen Mengen $V_1, V_2$ unter der stetigen Abbildung $f$, also sind $Z \cap f^{-1}(V_1), Z \cap f^{-1}(V_2) \in \mathcal{O}_Z$.
- $Z \cap f^{-1}(V_1), Z \cap f^{-1}(V_2)$ sind wegen $\emptyset \neq U_1, U_2$ und $U_1, U_2 \subseteq f(Z)$ nichtleer
- $(Z \cap f^{-1}(V_1)) \cap (Z \cap f^{-1}(V_2) \subseteq f^{-1}(U_1) \cap f^{-1}(U_2) = f^{-1}(U_1 \cap U_2) = f^{-1}(\emptyset) = \emptyset$

im Widerspruch dazu steht, daß $Z$ zusammenhängend ist. □

**Satz 13.5.4.** *Es sei $\emptyset \neq Z \subseteq \mathbb{R}$, dann sind äquivalent:*

*a) $Z$ ist zusammenhängend (bezüglich der Standardtopologie $\mathcal{O}_{\mathbb{R}}$).*

*b) $Z$ ist ein Intervall.*

*Beweis.*

a)⇒b) Es sei $Z$ zusammenhängend.

- Besteht $Z$ aus nur einem Punkt, so ist $Z$ ein Intervall.
- Ist $\mathbb{R}\backslash Z = \emptyset$, so ist $Z = \mathbb{R}$ ein Intervall.
- Es seien $a, b \in Z$ mit $a < b$ und $x \in ]a, b[$. Angenommen $x \in \mathbb{R}\backslash Z \neq \emptyset$, dann ist $Z \subseteq \mathbb{R}\backslash\{x\} = ]-\infty, x[\cup]x, \infty[$. Da $]-\infty, x[$ und $]x, \infty[$ offen bezüglich $\mathcal{O}_{\mathbb{R}}$ sind, ist $Z = (Z \cap ]-\infty, x[) \cup (Z \cap ]x, \infty[)$ eine Zerlegung von $Z$ in zwei disjunkte, nichtleere offene Mengen. Dies ist ein Widerspruch, denn $Z$ ist zusammenhängend. Das zeigt, daß für je zwei Punkte $a, b \in Z$, $a < b$ auch jedes $x \in [a, b]$ in $Z$ enthalten ist, dh. daß $Z$ ein Intervall ist.

b)⇒a) Es sei $Z$ ein Intervall. Angenommen $Z$ ist nicht zusammenhängend, dann gibt es offene Mengen $U, V \in \mathcal{O}_Z$ mit $\emptyset \neq U, V \subseteq Z$, $U \cap V = \emptyset$ und $U \cup V = Z$. Wähle $a \in U$ und $b \in V$ und setze $a < b$ (eventuell nach Umbenennen von $U$ und $V$) voraus. Dann ist $M := \{x \in V : a < x\} \neq \emptyset$, da $b \in M$ und $z := \inf(M)$ existiert nach Korollar 4.1.7, da etwa $a$ eine untere Schranke von $M$ ist. Das zeigt $a \leq z \leq b$ und da $Z$ ein Intervall ist, ist auch noch $z \in Z$. Betrachte nun die beiden Fälle:

- $z \in U$: Zu jedem $n \in \mathbb{N}$ gibt es nach Definition des Infimums $x_n \in V$ mit $z \leq x_n \leq z + \frac{1}{n}$, deshalb ist die Folge $(x_n)_{n\in\mathbb{N}}$ konvergent und $z = \lim\limits_{n\to\infty} x_n$, also $z \in \overline{V}$ gemäß Satz 13.2.9. Wegen $V = Z\backslash U$ ist $V$ (als Komplement der offenen Menge $U$) abgeschlossen (in $\mathcal{O}_Z$), also $V = \overline{V}$ (in $\mathcal{O}_Z$) und daher $z \in U \cap V \neq \emptyset$; Widerspruch.
- $z \in V$: Dann ist $z \neq a \in U$ und da $V \in \mathcal{O}_Z$ gibt es eine offene Umgebung $W$ von $z$ (bezüglich $\mathcal{O}_{\mathbb{R}}$) und $R > 0$ mit $]z - R, z + R[\subseteq W$ und $W \cap Z \subseteq V$. Wähle nun $r := \min\{R, |a - z|, |b - z|\}$ und da $Z$ ein Intervall ist, gilt dann $]z - r, z + r[\subseteq W \cap Z \subseteq V$. Wir können also ein $w \in V$ mit $a < w < z$ wählen, aber damit ist $w \in M$ und $w < z = \inf(M)$; Widerspruch.

□

**Satz 13.5.5** (Zwischenwertsatz). *Es sei $(X, \mathcal{O}_X)$ ein topologischer Raum, $Z \subseteq X$ zusammenhängend und $f : X \to \mathbb{R}$ sei stetig. Dann ist $f(Z)$ ein Intervall.*

*Beweis.* $f(Z)$ ist als Bild der zusammenhängenden Menge $Z$ unter der stetigen Abbildung $f$ wieder zusammenhängend, also als zusammenhängende Teilmenge von $\mathbb{R}$ ein Intervall. □

**Rezept 13.5.6** (Bisektionsverfahren)**.** Es sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \to \mathbb{R}$ stetig und $a, b \in I$ mit $a < b$ und $f(a)f(b) < 0$. Da $f(I)$ ein Intervall ist und $f(a)$ und $f(b)$ verschiedenenes Vorzeichen haben, gibt es mindestens eine Nullstelle $\xi_* \in ]a, b[$ von $f$, dh. $f(\xi_*) = 0$. Durch folgendes Iterationsverfahren läßt sich eine Folge $([x_n, y_n])_{n\in\mathbb{N}}$ von Intervallen angeben mit $\bigcap\limits_{n\in\mathbb{N}} [x_n, y_n] = \{\xi_*\}$ und $f(\xi_*) = 0$:

- Setze $x_0 := a$ und $y_0 := b$, dann gibt es $\xi_0 \in ]x_0, y_0[$ mit $f(\xi_0) = 0$.
- Für $n \geq 0$ setze

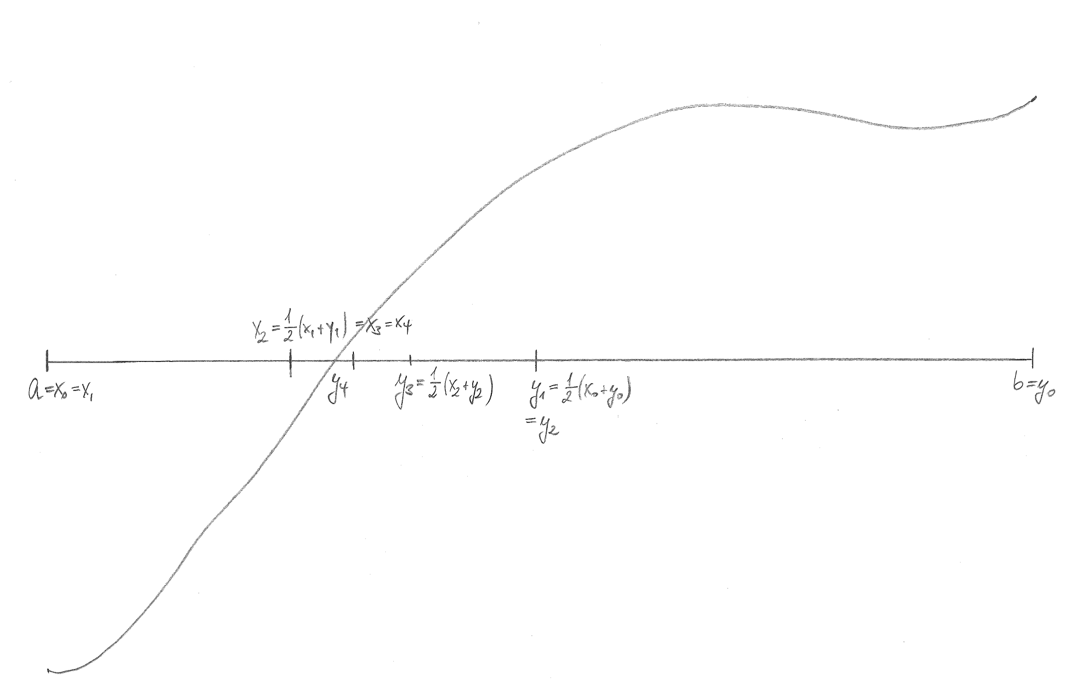
$$x_{n+1} := \begin{cases} x_n & \text{falls } f(x_n)f(\frac{1}{2}(x_n + y_n)) < 0 \\ \frac{1}{2}(x_n + y_n) & \text{falls } f(x_n)f(\frac{1}{2}(x_n + y_n)) \geq 0 \end{cases} \tag{13.5.1}$$

$$y_{n+1} := \begin{cases} \frac{1}{2}(x_n + y_n) & \text{falls } f(x_n)f(\frac{1}{2}(x_n + y_n)) \leq 0 \\ y_n & \text{falls } f(x_n)f(\frac{1}{2}(x_n + y_n)) \geq 0 \end{cases} \tag{13.5.2}$$

  dann gibt es $\xi_{n+1} \in ]x_{n+1}, y_{n+1}[$ mit $f(\xi_{n+1}) = 0$ falls $]x_{n+1}, y_{n+1}[ \neq \emptyset$ oder $f(\frac{1}{2}(x_n + y_n)) = 0$ falls $x_{n+1} = y_{n+1}$.

Nach Iteration ist im $n$-ten Schritt eine Nullstelle in $]x_n, y_n[$ bis auf einen Fehler von höchstens $|y_n - x_n| = 2^{-n}|y_0 - x_0|$ bestimmt und $\lim\limits_{n\to\infty} \xi_n = \xi_*$ erfüllt $f(\xi_*) = 0$.

**Bemerkung 13.5.7.** Die folgende Skizze zeigt, wie man sich beim Bisektionsverfahren durch Halbieren des Intervalls und durch Aussuchen des Teils, in dem ein Vorzeichenwechsel stattfindet an die Nullstelle herantastet:



**Lemma 13.5.8.** *Es sei $(X, \mathcal{O}_X)$ ein topologischer Raum, $I$ eine Menge und für alle $i \in I$ sei $Z_i \subseteq X$ zusammenhängend mit $\bigcap\limits_{i\in I} Z_i \neq \emptyset$. Dann ist $Z := \bigcup\limits_{i\in I} Z_i$ zusammenhängend.*

*Beweis.* Wähle $z \in \bigcap\limits_{i\in I} Z_i \neq \emptyset$ und es sei $\emptyset \neq W \subseteq \bigcup\limits_{i\in I} Z_i =: Z$ offen und abgeschlossen bezüglich $\mathcal{O}_Z$. Für jedes $j \in I$ mit $Z_j \cap W \neq \emptyset$ gilt, daß $\emptyset \neq Z_j \cap W$ offen und abgeschlossen bezüglich $\mathcal{O}_{Z_j}$ ist. Weil $Z_j$ zusammenhängend ist, folgt daraus $Z_j = Z_j \cap W$, dh. wir haben gezeigt:

$$Z_j \cap W \neq \emptyset \quad \Rightarrow \quad Z_j \subseteq W. \tag{13.5.3}$$

Weil $\emptyset \neq W \subseteq \bigcup\limits_{i\in I} Z_i$ ist, gibt es $j \in I$ mit $W \cap Z_j \neq \emptyset$, also ist $z \in Z_j \subseteq W$. Wegen $z \in \bigcap\limits_{i\in I} Z_i$ folgt also $z \in W \cap Z_i$ für alle $i \in I$, womit $Z_i \subseteq W$ für alle $i \in I$ aus (13.5.3) folgt. Damit ist $W = \bigcup\limits_{i\in I} Z_i$, also $\bigcup\limits_{i\in I} Z_i$ zusammenhängend. □

**Lemma 13.5.9.** *Ist $(X, \mathcal{O}_X)$ ein topologischer Raum und $x \in X$, dann heißt*

$$Z(x) := \bigcup_{\substack{Z \subseteq X \text{ zusammenhängend} \\ x \in Z}} Z \tag{13.5.4}$$

*die* **Zusammenhangskomponente** *von $x$. Dann ist $Z(x)$ die größte zusammenhängende Menge, die den Punkt $x$ enthält und die Mengen $Z(x); x \in X$ bilden eine Zerlegung von $X$.*

*Beweis.* $\{x\}$ ist zusammenhängend, daher wird in (13.5.4) nicht über $\emptyset$ vereinigt. Nach Lemma 13.5.8 ist $Z(x)$ zusammenhängend, da $x$ in jeder der zusammenhängenden Mengen über die die Vereinigung gebildet wird enthalten ist. Ebenso folgt dann daß die Zusammenhangskomponenten $Z(x), x \in X$ eine Zerlegung von $X$ bilden, denn da $x \in Z(x)$ ist, ergibt sich $X$ als Vereinigung aller Zusammenhangskomponenten und falls $z \in Z(x) \cap Z(y)$ ist, so ist $Z(x) \cap Z(y) \neq \emptyset$, also ist $Z(x) \cup Z(y)$ nach Lemma 13.5.8 zusammenhängend, also $Z(x) = Z(z) = Z(y)$. □

**Lemma 13.5.10.** *Jede nichtleere offene Teilmenge $U \subseteq \mathbb{R}$ (bezüglich der Standardtopologie) ist eine abzählbare Vereinigung von paarweise disjunkten offenen Intervallen.*

*Beweis.* Es sei $\emptyset \neq U \subseteq \mathbb{R}$ offen und zu $x \in U$, sei $Z(x)$ die Zusammenhangskomponente von $U$ in der $x$ enthalten ist. Dann ist $Z(x)$ eine offene Teilmenge von $\mathbb{R}$, denn da $U$ offen ist gibt es für jedes $y \in Z(x)$ ein $r = r(y, U) > 0$ mit $]y - r, y + r[ \subseteq U$. Da $y \in Z(x) \cap ]y - r, y + r[ \neq \emptyset$ ist $Z(x) \cup ]y - r, y + r[$ nach Lemma 13.5.8 zusammenhängend, also $]y - r, y + r[ \subseteq Z(x)$ und daher enthält $Z(x)$ mit jedem $y \in Z(x)$ auch ein offenes Intervall, dh. $Z(x)$ ist offene Teilmenge von $\mathbb{R}$. Als offene und zusammenhängende Teilmenge von $\mathbb{R}$ ist $Z(x)$ nach Satz 13.5.4 ein offenes Intervall und damit gibt es nach Lemma 13.5.9 eine Teilmenge $V \subseteq U$, so daß $U = \bigcup\limits_{x\in V} Z(x)$, $Z(x)$ ist ein nichtleeres offenes Intervall für jedes $x \in V$ und $Z(x) \cap Z(y) = \emptyset$ für alle $x, y \in V$ mit $x \neq y$. Da $\mathbb{Q}$ dicht in $\mathbb{R}$ ist, folgt $\mathbb{Q} \cap Z(x) \neq \emptyset$ für alle $x \in V$ nach Lemma 13.1.9 und daher folgt aus der Abzählbarkeit von $\mathbb{Q}$ die Abzählbarkeit von $V$. □

## 13.6 Kompaktheit

**Definition 13.6.1.** *Ein topologischer Raum $(X, \mathcal{O})$ heißt* **kompakt***, wenn*

1. *$X$ hausdorffsch*
2. *Zu jeder Familie $U_i, i \in I$ mit $U_i \in \mathcal{O}$ und $\bigcup\limits_{i \in I} U_i = X$ ("offene Überdeckung") gibt es eine endliche Menge $J \subseteq I$ mit $\bigcup\limits_{i \in J} U_i = X$ ("endliche Teilüberdeckung")*

*Eine Teilmenge $K \subseteq X$ des topologischen Raums $(X, \mathcal{O})$ ist*

- *kompakt, wenn $(K, \mathcal{O}_K)$ in der Relativtopologie $\mathcal{O}_K := \{U \cap K : U \in \mathcal{O}\}$ kompakt ist.*
- **relativ kompakt***, wenn der Abschluß $\overline{K}$ kompakt ist.*
- **relativ kompakt in** *$V \subseteq X$, wenn $K$ relativ kompakt und $\overline{K} \subseteq V$ ist.*

**Lemma 13.6.2.** *Es sei $(X, \mathcal{O})$ ein topologischer Raum und $K \subseteq X$ kompakt.*

*a) Ist $X$ hausdorffsch, so ist $K$ abgeschlossen.*

*b) Ist $A \subseteq K$ abgeschlossen, so ist $A$ kompakt.*

*Beweis.* a) Es sei ohne Einschränkung $X \backslash K \neq \emptyset$. Wähle $y \in X \backslash K$, dann gibt es zu jedem $x \in K$ offene Umgebungen $U_x$ von $x$ und $V_x$ von $y$ mit $U_x \cap V_x = \emptyset$. Weil $K = \bigcup\limits_{x \in K} (U_x \cap K)$ eine offene Überdeckung von $K$ (bezüglich $\mathcal{O}_K$) ist, gibt es wegen Kompaktheit von $K$ Punkte $x_1, ..., x_N \in K$ mit $K \subseteq U_{x_1} \cup ... \cup U_{x_N}$. Folglich ist $V := V_{x_1} \cap ... \cap V_{x_N}$ eine offene Umgebung von $y$ mit $V \cap K = \emptyset$; damit ist $X \backslash K$ offen, also $K$ abgeschlossen.

b) Sind $U_i \in \mathcal{O}$ und $A \subseteq \bigcup\limits_{i \in I} U_i$, dann ist $K = (K \backslash A) \cup \bigcup\limits_{i \in I} (U_i \cap K)$ eine offene Überdeckung von $K$ in der Relativtopologie. Damit gibt es eine endliche Menge $J \subseteq I$ mit $K = (K \backslash A) \cup \bigcup\limits_{i \in J} (U_i \cap K)$, daher ist $A \subseteq \bigcup\limits_{i \in J} U_i$, also $A$ kompakt.

□

**Lemma 13.6.3.** *Sei $X$ ein kompakter topologischer Raum, $Y$ Hausdorffraum und $f : X \to Y$ stetig, dann ist $f(X)$ kompakt.*

*Beweis.* Ist $(V_i)_{i \in I}$ eine offene Überdeckung von $f(X)$, dann ist $(f^{-1}(V_i))_{i \in I}$ eine offene Überdeckung von $X$. Da $X$ kompakt ist, gibt es also eine endliche Teilmenge $J \subseteq I$ mit $\bigcup\limits_{i \in J} f^{-1}(V_i) = X$ und folglich ist $f(X) \subseteq \bigcup\limits_{i \in J} V_i$. □

**Satz 13.6.4** (Tychonoff)**.** *Es sei $I$ eine Menge und für jedes $i \in I$ sei $(X_i, \mathcal{O}_{X_i})$ ein kompakter topologischer Raum. Dann ist auch $X := \prod\limits_{i \in I} X_i$ versehen mit der Produkttopologie ein kompakter topologischer Raum.*

*Beweis.* siehe Chernoff, American Mathematical Monthly Nummer 99, Seite 932-934 (1992) □

**Definition 13.6.5.** *Eine Teilmenge $Y \subseteq X$ eines metrischen Raums $(X, d)$ heißt* **totalbeschränkt***, wenn es zu jedem $\varepsilon > 0$ endlich viele Mengen $U_1, ..., U_N$ mit Durchmesser*

$$\delta(U_j) := \sup\{d(u, v) : u, v \in U_j\} < \varepsilon \tag{13.6.1}$$

*für $j = 1, ..., N$ und $Y \subseteq U_1 \cup ... \cup U_N$ gibt.*

**Bemerkung 13.6.6.** Jede totalbeschränkte Menge ist auch beschränkt, aber die Umkehrung gilt im Allgemeinen nicht, zB.

a) Ist $(X, d)$ ein metrischer Raum, so gibt es ganz einfache Beispiele von Mengen, die beschränkt aber nicht totalbeschränkt sind. Etwa $\mathbb{R}$ versehen mit der Metrik gegeben durch $d(x, y) := \begin{cases} 1 & \text{falls } x \neq y \\ 0 & \text{falls } x = y \end{cases}$ ist eine beschränkte Menge, zB. $\mathbb{R} = \{y \in \mathbb{R} : d(0, y) < 2\}$ aber nicht totalbeschränkt denn zB. sind alle offenen Kugeln $K(x, \frac{1}{2}) = \{y \in \mathbb{R} : d(x, y) < \frac{1}{2}\} = \{x\}$ paarweise disjunkt und damit hat $\mathbb{R}$ keine endliche Überdeckung durch Kugeln mit Radius $\frac{1}{2}$.

b) Im vollständigen Raum $l^2(\mathbb{N})$ ist

$$\overline{K}(0, 1) = \{x \in l^2(\mathbb{N}) : \|x\| \leq 1\}$$

beschränkt, aber nicht totalbeschränkt, denn $\frac{1}{2}e_k = (0, ..., 0, \frac{1}{2}, 0, ...) \in \overline{K}(0, 1) \subseteq l^2(\mathbb{N})$ und für alle $k, l \in \mathbb{N}$, $k \neq l$ gilt

$$\|\frac{1}{2}e_k - \frac{1}{2}e_l\| = \frac{\sqrt{2}}{2}. \tag{13.6.2}$$

Ist also $\varepsilon < \frac{\sqrt{2}}{4}$, so bilden $\overline{K}(\frac{1}{2}e_k, \varepsilon), k \in \mathbb{N}$ eine Folge von paarweise disjunkten Mengen in $\overline{K}(0, 1)$, also $\overline{K}(0, 1)$ nicht totalbeschränkt. $\overline{K}(0, 1)$ ist als abgeschlossene Kugel abgeschlossen und als Teilmenge des vollständigen Raums $l^2(\mathbb{N})$ wieder vollständig, $\overline{K}(0, 1)$ ist nicht kompakt, denn etwa $(\frac{1}{2}e_k)_{k\in\mathbb{N}}$ hat wegen (13.6.2) keine konvergente Teilfolge. Das zeigt, daß man auf die Totalbeschränktheit im Allgemeinen nicht verzichten kann.

**Satz 13.6.7.** *Es sei $(X, d)$ ein metrischer Raum, $Y \subseteq X$, dann sind äquivalent:*

*a) $Y$ kompakt.*

*b) Jede Folge $(x_n)_{n\in\mathbb{N}}$ in $Y$ hat mindestens eine Teilfolge, die in $Y$ konvergiert.*

*c) $Y$ totalbeschränkt und vollständig.*

*Beweis.*

*a)* $\Rightarrow$ *b)* Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $Y$, $Y \subseteq X$ kompakt und $A_n := \overline{\{x_n, x_{n+1}, ..., \}}$ der Abschluß von $\{x_n, x_{n+1}, ..., \}$. Angenommen

$$\bigcap_{n\in\mathbb{N}} A_n = \emptyset,$$

dann bilden die in der Relativtopologie von $Y$ offenen Mengen $U_n := (X\backslash A_n) \cap Y$ eine offene Überdeckung von $Y$:

$$\bigcup_{n\in\mathbb{N}} U_n = Y \cap \left(\bigcup_{n\in\mathbb{N}} (X\backslash A_n)\right) = Y \cap \left(X\backslash \bigcap_{n\in\mathbb{N}} A_n\right) = Y \cap X = Y$$

Weil $Y$ kompakt ist, gibt es daher endlich viele $U_1, ..., U_N$ mit $Y = U_1 \cup ... \cup U_N$. Dies bedeutet aber $Y \cap A_1 \cap .. \cap A_N = \emptyset$, was aber im Widerspruch dazu steht, daß für $n > N$ der Punkt $x_n \in Y \cap A_1 \cap ... \cap A_N$ ist. Wir haben damit gezeigt, daß

$$\bigcap_{n\in\mathbb{N}} \overline{\{x_n, x_{n+1}, ...\}} \neq \emptyset, \tag{13.6.3}$$

womit wir also gezeigt haben, daß die Folge $(x_n)_{n\in\mathbb{N}}$ in $Y$ wenigstens eine konvergente Teilfolge besitzt.

*b)* $\Rightarrow$ *c)* Nach Voraussetzung b) hat jede Cauchyfolge eine konvergente Teilfolge und ist damit konvergent. Somit ist $Y$ vollständig. Angenommen $Y$ ist nicht totalbeschränkt, so gibt es ein $\varepsilon_0 > 0$, und eine Folge $(x_n)_{n\in\mathbb{N}}$ in $Y$, so daß $x_{N+1} \in Y\backslash \bigcup_{n=1}^{N} K(x_n, \varepsilon_0) \neq \emptyset$ für jedes $N \in \mathbb{N}$ gilt. Insbesondere ist $d(x_n, x_m) > \varepsilon_0$ falls $n \neq m$, also hat $(x_n)_{n\in\mathbb{N}}$ keine konvergente Teilfolge.

*c)* $\Rightarrow$ *a)* Angenommen $(U_i)_{i\in I}$ ist eine offene Überdeckung von $Y$, zu der es keine endliche Teilüberdeckung gibt. Wir definieren nun induktiv eine Folge von Kugeln $K_n$ mit Radius $2^{-n}$ und Mittelpunkten $x_n \in Y$ wie folgt: $K_1 = K(x_1, 1)$ sei eine Kugel vom Radius 1, die nicht durch endlich viele Elemente $U_i$ überdeckt werden kann. Da $Y$ totalbschränkt ist und $(U_i)_{i\in I}$ keine endliche Teilüberdeckung hat, muß es eine solche Kugel $K_1$ geben. Es sei nun $K_{n-1} = K(x_{n-1}, 2^{1-n})$ für $n \geq 2$ so gewählt, daß es keine endliche Überdeckung von $K_{n-1}$ in $(U_i)_{i\in I}$ gibt. Wir betrachten nun eine endliche Überdeckung von $Y$ durch Kugeln mit Radius $2^{-n}$, dann gibt es unter den Kugeln, die mit $K_{n-1}$ nichtleeren Durchschnitt haben, mindestens eine Kugel $K_n = K(x_n, 2^{-n})$, die nicht durch endlich viele Elemente aus $(U_i)_{i\in I}$ überdeckt werden kann. In der so konstruierten Folge $(K(x_n, 2^{-n}))_{n\in\mathbb{N}}$ haben zwei Kugeln $K(x_{n-1}, 2^{1-n})$ und $K(x_n, 2^{-n})$ immer nichtleeren Durchschnitt; mit $y \in K(x_{n-1}, 2^{1-n}) \cap K(x_n, 2^{-n})$ gilt nach Dreiecksungleichung:

$$d(x_{n-1}, x_n) \leq d(x_{n-1}, y) + d(y, x_n) \leq 2^{1-n} + 2^{-n} < 2^{2-n},$$

daher ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $Y$, die wegen der Vollständigkeit gegen ein $x \in Y$ konvergiert. Es sei $i_0 \in I$, so daß $x \in U_{i_0}$. Da $U_{i_0}$ offen ist, gibt es ein $r > 0$ mit $K(x, r) \subseteq U_{i_0}$. Wegen $x = \lim_{n\to\infty} x_n$ existiert ein $m \in \mathbb{N}$ mit $d(x, x_m) < \frac{r}{2}$ und $2^{-m} < \frac{r}{2}$; aus der Dreiecksungleichung folgt dann $K(x_m, 2^{-m}) \subseteq K(x, r) \subseteq U_{i_0}$ im Widerspruch zur Konstruktion von $K(x_m, 2^{-m})$.

□

**Lemma 13.6.8.** *Sei $X$ ein topologischer Raum, $K \subseteq X$ sei kompakt und $g : X \to \mathbb{R}$ stetig. Dann existieren $x_{\min}, x_{\max} \in K$ mit $g(x_{\max}) = \max_{x\in K} g(x) < \infty$ und $g(x_{\min}) = \min_{x\in K} g(x) > -\infty$.*

*Beweis.* Nach Lemma 13.6.3 ist $g(K) \subseteq \mathbb{R}$ kompakt, also (total-)beschränkt (nach Satz 13.6.7) und abgeschlossen (nach Lemma 13.6.2). Somit existieren $\inf(g(K))$ und $\sup(g(K))$ nach Satz 4.1.4 bzw. Lemma 4.1.6 und es gibt Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ in $g(K)$ mit $\lim_{n\to\infty} x_n = \inf(g(K))$ und $\lim_{n\to\infty} y_n = \sup(g(K))$. Da $g(K)$ abgeschlossen ist, folgt daher $\inf(g(K)) \in \overline{g(K)} = g(K)$ und $\sup(g(K)) \in \overline{g(K)} = g(K)$, also ist $\inf(g(K)) = \min(g(K)) \in g(K)$ und $\sup(g(K)) = \max(g(K)) \in g(K)$. □

**Satz 13.6.9 (Bolzano-Weierstraß).** [7] *Sei $d \in \mathbb{N}$ und $A \subseteq \mathbb{R}^d$, dann sind äquivalent:*

*a) $A$ ist kompakt (in der von $\|\cdot\|_\infty$ definierten Normtopologie $\mathcal{O}_{\|\cdot\|_\infty}$).*

*b) $A$ ist beschränkt (bezüglich $\|\cdot\|_\infty$) und abgeschlossen (bezüglich $\mathcal{O}_{\|\cdot\|_\infty}$).*

*Beweis.* *a)* $\Rightarrow$ *b)*: Nach Satz 13.6.7 ist $A$ totalbeschränkt, also insbesondere beschränkt und nach Lemma 13.6.2 ist $A$ abgeschlossen.
*b)* $\Rightarrow$ *a)*: Da $A$ in $\|\cdot\|_\infty$ beschränkt ist, gibt es $x = (x_1, ..., x_d) \in \mathbb{R}^d$ und $r \in ]0, \infty[$ mit

$$\overline{K}(x, r) = [x_1 - r, x_1 + r] \times ... \times [x_d - r, x_d + r] \supseteq A.$$

Da $A$ abgeschlossen ist, genügt es nach Lemma 13.6.2 die Kompaktheit von $\overline{K}(x, r)$ nachzuweisen. Sei $(y_n)_{n\in\mathbb{N}} = (y_{n,1}, ..., y_{n,d})_{n\in\mathbb{N}}$ eine Folge in $\overline{K}(x, r)$, dann gibt es nach Lemma 5.5.1 eine monotone Teilfolge $(y_{n_1(k),1})_{k\in\mathbb{N}}$ der reellen Folge $(y_{n,1})$ und wegen $y_{n,1} \in [x_1 - r, x_1 + r]$ konvergiert die beschränkte und monotone Folge $y_{n_1(k),1} \xrightarrow{k\to\infty} y^{(1)}$. Wähle nun analog für $j = 1, ..., d-1$ eine Teilfolge $(n_{j+1}(k))_{k\in\mathbb{N}}$ von $(n_j(k))_{k\in\mathbb{N}}$, so daß $(y_{n_{j+1}(k),j+1})_{k\in\mathbb{N}}$ eine monotone Teilfolge von $(y_{n_j(k),j+1})_{k\in\mathbb{N}}$ ist, also

$$y_{n_{j+1}(k),j+1} \xrightarrow{k\to\infty} y^{(j+1)} \in [x_{j+1} - r, x_{j+1} + r]$$

erfüllt, dann gilt $\lim_{k\to\infty} y_{n_d(k)} = (y^{(1)}, ..., y^{(d)})$ in der $\|\cdot\|_\infty$ Norm, damit ist $\overline{K}(x, r)$ nach Satz 13.6.7 kompakt. □

**Satz 13.6.10.** *Seien $X, Y$ metrische Räume, $X$ kompakt und $f : X \to Y$ stetig, dann ist $f$ gleichmäßig stetig.*

*Beweis.* Wenn $f$ nicht gleichmäßig stetig ist, gibt es ein $\varepsilon > 0$ und Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ in $X$ mit $d(x_n, y_n) \leq \frac{1}{n}$ und $d'(f(x_n), f(y_n)) > \varepsilon$. Da $X$ kompakt ist, gibt es konvergente Teilfolgen $(x_{n_k})_{k\in\mathbb{N}}$ und $(y_{n_k})_{k\in\mathbb{N}}$ und $x := \lim_{k\to\infty} x_{n_k}$ und $y := \lim_{k\to\infty} y_{n_k}$ existieren. Dann ist aber $x = y$ und aufgrund der Stetigkeit von $f$ und $d'$ gilt $\lim_{k\to\infty} d'(f(x_{n_k}), f(y_{n_k})) = d'(f(x), f(y)) = 0$ im Widerspruch zur Wahl von $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ und $\varepsilon$. □

[7] wird manchmal auch als Satz von Heine-Borel bezeichnet

**Satz 13.6.11.** *Ist $X$ ein kompakter metrischer Raum, $(Y,d)$ ein vollständiger metrischer Raum, so ist*

$$C(X,Y) := \{f : X \to Y : f \text{ ist stetig } \} \tag{13.6.4}$$

*versehen mit*

$$d_\infty(f,g) := \sup\{d(f(x),g(x)) : x \in X\} \tag{13.6.5}$$

*ein vollständiger metrischer Raum.*

*Beweis.* Für $f,g \in C(X,Y)$ ist

$$\begin{array}{rcl} \phi : X & \to & Y \times Y & \to & [0,\infty[ \\ x & \mapsto & (f(x),g(x)) & \mapsto & d(f(x),g(x)) \end{array}$$

– wobei $Y \times Y$ mit der Produkttopologie versehen wird – als Komposition stetiger Abbildungen und wegen Lemma 13.1.20 wieder stetig, daher nimmt $\phi$ auf der kompakten Menge $X$ ein Maximum an und somit ist $d_\infty(f,g) \in [0,\infty[$ wohldefiniert. Die Eigenschaften einer Metrik folgen dann aus denen von $d$. Ist $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $(C(X,Y),d_\infty)$, so gibt es für alle $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$ mit

$$d_\infty(f_m,f_n) = \sup\{d(f_m(x),f_n(x)) : x \in X\} \le \varepsilon \qquad \text{für alle } m,n \ge N(\varepsilon). \tag{13.6.6}$$

Für jedes $x \in X$ ist also $(f_n(x))_{n\in\mathbb{N}}$ eine Cauchyfolge in $Y$, die aufgrund der Vollständigkeit von $Y$ den Grenzwert $f(x) := \lim\limits_{n\to\infty} f_n(x)$ besitzt; daher ist

$$\begin{array}{rcl} f : X & \to & Y \\ x & \mapsto & f(x) \end{array}$$

definiert. Aufgrund dieser punktweisen Konvergenz gibt es zu jedem $\varepsilon > 0$ und $x \in X$ ein $N(x,\varepsilon) \in \mathbb{N}$ mit $d(f(x),f_n(x)) < \varepsilon$ für $n \ge N(x,\varepsilon)$. Mit $L(x,\varepsilon) := \max\{N(\varepsilon),N(x,\varepsilon)\}$ gilt:

$$\sup\{d(f(x),f_n(x)) : x \in X\} \le \sup\{d(f(x),f_{L(x,\varepsilon)}(x)) + d(f_{L(x,\varepsilon)}(x),f_n(x)) : x \in X\} \le 2\varepsilon$$

für alle $n \ge N(\varepsilon)$. Wegen der Stetigkeit von $f_N$ gibt es eine Umgebung $U$ von $a$ mit $d(f_N(x),f_N(a)) < \varepsilon$ für alle $x \in U$, dann gilt:

$$d(f(a),f(x)) \le d(f(a),f_N(a)) + d(f_N(a),f_N(x)) + d(f_N(x),f(x)) < 5\varepsilon$$

für alle $x \in U$, $N \ge N(\varepsilon)$ und damit ist $f$ stetig im Punkt $a \in X$ und weil $a$ beliebig war auf ganz $X$. Damit ist $f \in C(X,Y)$ und $(f_n)_{n\in\mathbb{N}}$ konvergiert in der Metrik $d_\infty$ gegen $f$. □

**Satz 13.6.12 (Arzela-Ascoli).** *Es sei $(X,d)$ ein kompakter metrischer Raum und $(Y,d')$ ein vollständiger metrischer Raum. Eine Menge $M \subseteq C(X,Y)$ ist genau dann kompakt in der durch $d_\infty$ gegebenen Topologie, wenn*

*a) Für alle $x \in X$ die Menge $M(x) := \{f(x) : f \in M\} \subseteq Y$ relativ kompakt ist und*

b) *$M$* **gleichgradig stetig***: Für alle $\varepsilon > 0$ gibt es $\delta > 0$, so daß für alle $f \in M$ und alle $x, y \in X$ mit $d(x,y) < \delta$ gilt: $d'(f(x), f(y)) < \varepsilon$.*

c) *$M$ abgeschlossen.*

*Beweis.* "$\Rightarrow$": Ist $M \subseteq C(X,Y)$ kompakt, so ist $M$ nach Lemma 13.6.2 abgeschlossen. Wegen $d'(f(x), g(x)) \leq d_\infty(f,g)$ ist $\begin{array}{rcl} \phi_x : C(X,Y) & \to & Y \\ f & \mapsto & f(x) \end{array}$ stetig, also $M(x) = \phi_x(M)$ nach Lemma 13.6.3 kompakt. Zu $\varepsilon > 0$ wähle $f_1, ..., f_N \in C(X,Y)$ mit

$$M \subseteq \bigcup_{j=1}^{N} \{g \in C(X,Y) : d_\infty(g, f_j) < \frac{\varepsilon}{3}\}.$$

Die stetigen Funktionen $f_1, ..., f_N$ sind nach Satz 13.6.10 gleichmäßig stetig, daher gibt es $\delta > 0$, so daß $d'(f_j(x), f_j(y)) < \frac{\varepsilon}{3}$ für $j = 1, ..., N$ und alle $x, y \in X$ mit $d(x,y) < \delta$. Dann ist aber

$$\begin{aligned} d'(f(x), f(y)) &\leq d'(f(x), f_j(x)) + d'(f_j(x), f_j(y)) + d'(f_j(y), f(y)) \leq \\ &\leq 2d_\infty(f, f_j) + d'(f_j(x), f_j(y)) < \varepsilon, \end{aligned}$$

wenn $d_\infty(f, f_j) < \frac{\varepsilon}{3}$ und $d(x,y) < \delta$. Damit ist $M$ gleichgradig stetig.
"$\Leftarrow$": Es sei $M \subseteq C(X,Y)$ abgeschlossen, dann ist $M$ als abgeschlossene Teilmenge des nach Satz 13.6.11 vollständigen $C(X,Y)$ wieder vollständig. Es sei $\varepsilon > 0$; weil $M$ gleichgradig stetig ist, gibt es ein $\delta > 0$, so daß $d'(f(x), f(y)) < \frac{\varepsilon}{4}$ für alle $f \in M$ erfüllt ist, solange $d(x,y) < \delta$. Da $X$ kompakt ist, gibt es $x_1, ..., x_N \in X$ mit $X = \bigcup_{j=1}^{N} \{x \in X : d(x_j, x) < \delta\}$. Da alle $M(x_j)$ relativ kompakt sind, gibt es $y_1, ..., y_P \in Y$ mit

$$B := \bigcup_{j=1}^{N} M(x_j) \subseteq \bigcup_{k=1}^{P} \{y \in Y : d'(y, y_k) < \frac{\varepsilon}{4}\}.$$

Zu $\varphi \in \Phi := \{\varphi : \{1, ..., N\} \to \{1, ..., P\}\}$ definiere

$$M_\varphi := \{f \in M : d'(f(x_j), y_{\varphi(j)}) \leq \frac{\varepsilon}{4} \text{ für alle } j = 1, ..., N\}$$

dann ist $M \subseteq \bigcup_{\varphi \in \Phi} M_\varphi$. Ist nun $\varphi \in \Phi$ und $f, g \in M_\varphi$, so gibt es zu jedem $x \in X$ ein $j \in \{1, ..., N\}$ mit $d(x, x_j) < \delta$. Daraus folgt $d'(f(x), f(x_j)), d'(g(x), g(x_j)) \leq \frac{\varepsilon}{4}$ und somit

$$d'(f(x), g(x)) \leq d'(f(x), f(x_j)) + d'(f(x_j), y_{\varphi(j)}) + d'(y_{\varphi(j)}, g(x_j)) + d'(g(x_j), g(x)) \leq \varepsilon$$

$M_\varphi$ hat also Durchmesser $\leq \varepsilon$ und somit ist $M$ totalbeschränkt. Als vollständige, totalbeschränkte Menge ist $M$ nach Satz 13.6.7 kompakt. $\square$

## 13.7 Normierte Räume: Äquivalente Normen

**Definition 13.7.1.** *Es sei $V$ ein $\mathbb{K}$-Vektorraum. Zwei Normen $\|\cdot\|$ und $|||\cdot|||$ auf $V$ heißen* **äquivalente Normen**, *wenn es $0 < m \le M < \infty$ gibt, so daß*

$$m\|v\| \le |||v||| \le M\|v\| \tag{13.7.1}$$

*für alle $v \in V$ gilt.*

**Bemerkung 13.7.2.** Wie der Name dies schon suggeriert, definiert das eine Äquivalenzrelation auf der Menge der Normen von $V$.

**Beispiel 13.7.3.** Für jedes $x = (x_1, ..., x_d) \in \mathbb{K}^d$ gilt:

$$\frac{1}{d}\|x\|_1 = \frac{|x_1| + ... + |x_d|}{d} \le \|x\|_\infty = \max_{j=1,...,d} |x_j| \le \|x\|_1, \tag{13.7.2}$$

daher sind $\|\cdot\|_1$ und $\|\cdot\|_\infty$ äquivalente Normen auf $\mathbb{K}^d$.

**Lemma 13.7.4.** *Sind $\|\cdot\|$ und $|||\cdot|||$ Normen auf $V$, so sind äquivalent:*

*a)* *$\|\cdot\|$ und $|||\cdot|||$ sind äquivalente Normen.*

*b)* *Eine Folge $(x_n)_{n\in\mathbb{N}}$ konvergiert genau dann in $\|\cdot\|$, wenn $(x_n)_{n\in\mathbb{N}}$ in $|||\cdot|||$ konvergiert. Die Limiten stimmen überein.*

*c)* *Eine Folge $(x_n)_{n\in\mathbb{N}}$ konvergiert in $\|\cdot\|$ gegen $0$ genau dann, wenn $(x_n)_{n\in\mathbb{N}}$ in $|||\cdot|||$ gegen $0$ konvergiert.*

*Sind $\|\cdot\|$ und $|||\cdot|||$ äquivalente Normen auf $V$, so ist $(V, \|\cdot\|)$ genau dann vollständig, wenn $(V, |||\cdot|||)$ vollständig ist.*

*Beweis.* a) $\Rightarrow$ b) $\Rightarrow$ c) sind klar. *c)* $\Rightarrow$ *a)* Angenommen es gibt kein $M \in ]0, \infty[$, so daß die Ungleichung $|||v||| \le M\|v\|$ für alle $v \in V$ erfüllt ist, dann gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ in $V$ mit $|||x_n||| > n\|x_n\|$. Setze $y_n := \frac{x_n}{n\|x_n\|}$, so ist $\|y_n\| = \frac{1}{n} \to 0$, aber gleichzeitig ist $|||y_n||| > 1$, also c) nicht erfüllt. Analog geht es für die andere Abschätzung. Sind nun $\|\cdot\|$ und $|||\cdot|||$ äquivalente Normen, dann ist $(x_n)_{n\in\mathbb{N}}$ genau dann Cauchyfolge bzgl. $\|\cdot\|$, wenn es Cauchyfolge bzgl. $|||\cdot|||$ ist und wegen obiger Konvergenzüberlegungen ist $(x_n)_{n\in\mathbb{N}}$ genau dann konvergent bzgl. $\|\cdot\|$, wenn es bzgl. $|||\cdot|||$ konvergiert. Damit ist in diesem Fall $(V, \|\cdot\|)$ genau dann vollständig, wenn $(V, |||\cdot|||)$ vollständig ist. □

**Lemma 13.7.5.** *Es seien $\|\cdot\|$ und $|||\cdot|||$ äquivalente Normen auf $V$, dann ist die durch $\|\cdot\|$ definierte Topologie $\mathcal{O}_{\|\cdot\|}$ gleich der von $|||\cdot|||$ definierten Topologie $\mathcal{O}_{|||\cdot|||}$.*

*Beweis.* Ist $\emptyset \ne U \in \mathcal{O}_{\|\cdot\|}$, dann gibt es zu jedem $x \in U$ ein $r = r(x, U) > 0$ mit $K_{\|\cdot\|}(x, r) = \{v \in V : \|x - v\| < r\} \subseteq U$. Da $\|\cdot\|$ und $|||\cdot|||$ äquivalent sind gilt (13.7.1) und daher folgt $\|x - v\| \le \frac{1}{m}|||x - v|||$, also $K_{|||\cdot|||}(x, mr) \subseteq K_{\|\cdot\|}(x, r) \subseteq U$, dh. $U \in \mathcal{O}_{|||\cdot|||}$ oder $\mathcal{O}_{\|\cdot\|} \subseteq \mathcal{O}_{|||\cdot|||}$. Die andere Inklusion zeigt man ganz analog. □

**Beispiel* 13.7.6.** Auf $C([0,1])$ sind die beiden Normen $\|\cdot\|_\infty$ und $\|\cdot\|_1$ gegeben durch

$$\|f\|_\infty := \sup_{x\in[0,1]} |f(x)|$$
$$\|f\|_1 := \int\limits_0^1 |f(x)|dx$$

nicht äquivalent: $(f_n)_{n\in\mathbb{N}}$ mit $f_n(x) = x^n$ ist eine $\|\cdot\|_1$-Nullfolge: $\int\limits_0^1 x^n dx = \frac{1}{n+1} \to 0$. Und gleichzeitig ist $(f_n)_{n\in\mathbb{N}}$ keine $\|\cdot\|_\infty$-Nullfolge, denn $\|f_n\|_\infty = \sup\limits_{x\in[0,1]} |x^n| = 1$ für jedes $n \in \mathbb{N}$.

**Satz 13.7.7.** *a) Je zwei Normen auf $\mathbb{K}^d$ sind äquivalent.*

*b) Für $d \geq 1$ ist $(\mathbb{K}^d, \|\cdot\|)$ mit jeder Norm $\|\cdot\|$ vollständig.*

*c) $A \subseteq \mathbb{K}^d$ ist in der Normtopologie genau dann kompakt, wenn $A$ beschränkt und abgeschlossen ist.*

*Beweis.* a) Da jede Norm auf $\mathbb{C}^d$ auch eine Norm auf $\mathbb{R}^{2d}$ ist, dürfen wir uns auf den Fall $\mathbb{K} = \mathbb{R}$ beschränken. Es sei $x \in \mathbb{R}^d$, $x \neq 0$ und $e_1, ..., e_d$ die Standardbasis von $\mathbb{R}^d$ und $\|\cdot\|$ eine Norm auf $\mathbb{R}^d$. Dann ist

$$0 < C := \max\{\|e_1\|, ..., \|e_d\|\} < \infty.$$

Aus der Dreiecksungleichung und Beispiel 13.7.3 folgt

$$\|x\| = \|x_1e_1 + ... + x_de_d\| \leq \sum_{j=1}^d |x_j|\,\|e_j\| \leq C\sum_{j=1}^d |x_j| = C\|x\|_1 \leq Cd\|x\|_\infty. \quad (13.7.3)$$

Die Würfeloberfläche

$$S := \{x \in \mathbb{K}^d : \|x\|_\infty = 1\} = \|\cdot\|_\infty^{-1}(\{1\})$$

ist als Urbild der abgeschlossenen Menge $\{1\}$ bei der bezüglich $\mathcal{O}_{\|\cdot\|_\infty}$ stetigen Abbildung $\|\cdot\|_\infty : \mathbb{R}^d \to [0,\infty[$ wieder abgeschlossen und beschränkt, also nach Satz 13.6.9 kompakt. Wegen (13.7.3) ist die Funktion $\|\cdot\| : \mathbb{R}^d \to [0,\infty[$ stetig bezüglich $\mathcal{O}_{\|\cdot\|_\infty}$, daher nimmt $\|\cdot\|$ nach Lemma 13.6.8 auf $S$ ein Minimum $c \geq 0$ in einem Punkt $0 \neq y \in S$ an. Weil $\|\cdot\|$ eine Norm ist, folgt $\|y\| = c > 0$ aus der positiven Definitheit der Norm. Für alle $x \in \mathbb{K}^d$, $x \neq 0$ folgt nun aus $c = \min\limits_{\|x\|_\infty=1} \|x\| = \min\limits_{x\neq 0} \dfrac{\|x\|}{\|x\|_\infty}$ und (13.7.3) die Abschätzung

$$c\|x\|_\infty \leq \|x\| \leq Cd\|x\|_\infty. \quad (13.7.4)$$

Daher ist jede Norm $\|\cdot\|$ auf $\mathbb{R}^d$ zu $\|\cdot\|_\infty$ äquivalent und damit sind auch je zwei Normen zueinander äquivalent.

b) $(\mathbb{K}^d, \|\cdot\|_\infty)$ ist vollständig. Weil $\|\cdot\|$ und $\|\cdot\|_\infty$ nach a) äquivalente Normen sind, ist $(\mathbb{K}^d, \|\cdot\|)$ nach Lemma 13.7.4 vollständig.

c) Nach Teil a) sind alle Normen auf $\mathbb{K}^d$ äquivalent, daher können wir einfach von beschränkten Mengen sprechen ohne die Norm noch genauer zu spezifizieren. Nach Lemma 13.7.5 und Teil a) gibt es dann auch nur eine Normtopologie auf $\mathbb{K}^d$ und daher ist die Behauptung in Satz 13.6.9 bewiesen.

$\square$

## 13.8 Anwendung: Reelle Exponentialfunktion und natürlicher Logarithmus, $\pi$

In Satz 6.6.3 haben wir bereits die Funktionalgleichung $e^{w+z} = e^w e^z$ für die komplexe Exponentialfunktion gezeigt.

**Lemma 13.8.1.** *Die Exponentialfunktion* $\begin{array}{rcl} \exp : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & e^z \end{array}$ *ist stetig.*

*Beweis.* Nach Lemma 6.6.1 konvergiert die Exponentialreihe $e^z = \displaystyle\sum_{k=0}^{\infty} \frac{z^k}{k!}$ für alle $z \in \mathbb{C}$, daher definiert exp nach Satz 13.4.4 eine stetige Funktion $\exp : \mathbb{C} \to \mathbb{C}$. $\square$

**Lemma 13.8.2.** $\begin{array}{rcl} \exp_{\mathbb{R}} : \mathbb{R} & \to & ]0,\infty[ \\ x & \mapsto & e^x \end{array}$ *ist streng monoton steigend und bijektiv und in* $\widehat{\mathbb{R}}$ *gilt:*

$$\lim_{x\to\infty} (\exp|_{\mathbb{R}})(x) = \infty \qquad (13.8.1)$$

$$\lim_{x\to-\infty} (\exp|_{\mathbb{R}})(x) = 0 \qquad (13.8.2)$$

*Beweis.* Ist $h > 0$, so gilt nach Satz 6.6.3 und Einsetzen der Exponentialreihe

$$\frac{e^{x+h}}{e^x} = e^h = \sum_{k=0}^{\infty} \frac{h^k}{k!} > 1,$$

dh. für alle $x \in \mathbb{R}$ und $h > 0$ ist $e^{x+h} > e^x$, also ist $\begin{array}{rcl} \exp : \mathbb{R} & \to & ]0,\infty[ \\ x & \mapsto & e^x \end{array}$ streng monoton steigend und damit injektiv. Zum Nachweis der Surjektivität sei $y \in ]0,\infty[$.

- $y \geq 1$: Da $1 = e^0 \leq y < \displaystyle\sum_{k=0}^{\infty} \frac{y^k}{k!} = e^y$ und exp stetig und streng monoton steigend ist, folgt $\exp([0,y]) = [e^0, e^y] \ni y$ nach dem Zwischenwertsatz und damit gibt es ein $x \in [0,y]$ mit $e^x = y$.

- $y \in ]0,1[$, dann gibt es $x \in [0, \frac{1}{y}]$ mit $e^{-x} = y$, denn $\dfrac{1}{y} > 1$ und daher gibt es nach dem ersten Fall ein $x \in [0, \frac{1}{y}]$ mit $e^x = \frac{1}{y}$.

Daher ist $\exp|_{\mathbb{R}} : \mathbb{R} \to ]0,\infty[$ surjektiv und damit bijektiv. Da

$$\begin{array}{rcl} F : \widehat{\mathbb{R}} & \to & [-1,1] \\ x & \mapsto & \begin{cases} \dfrac{x}{1+|x|} & \text{für } x \in \mathbb{R} \\ -1 & \text{für } x = -\infty \\ 1 & \text{für } x = \infty \end{cases} \end{array}$$

streng monoton steigend ist, enthält jede Umgebung $U$ von $\infty$ (bzw. $-\infty$) bezüglich der von der Metrik $\begin{array}{rcl} d : \widehat{\mathbb{R}} \times \widehat{\mathbb{R}} & \to & [0,\infty[ \\ (x,y) & \mapsto & |F(x) - F(y)| \end{array}$ definierten Topologie $\mathcal{O}_d$ eine Menge $\{x \in \mathbb{R} : x > R\}$ (bzw. $\{x \in \mathbb{R} : x < -R\}$) und damit sind $\infty$ (und $-\infty$) Berührpunkte von $\mathbb{R}$. Für $x > 0$ ist

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \sum_{k=2}^{\infty} \frac{x^k}{k!} > x$$

also

$$d(e^x, \infty) = |F(e^x) - 1| \leq |F(x) - 1| = \left| \frac{x}{1+x} - 1 \right| = \frac{1}{1+x} = d(x,\infty)$$

und damit $\lim\limits_{x\to\infty} (\exp|_{\mathbb{R}})(x) = \infty$. Ebenso ist für $x < 0$

$$d(e^x, 0) = |F(e^x)| = \frac{e^x}{1+e^x} = \frac{1}{1+e^{|x|}} \leq \frac{1}{1+|x|} = d(x,-\infty)$$

und damit $\lim\limits_{x\to-\infty} (\exp|_{\mathbb{R}})(x) = 0$. □

**Definition 13.8.3.** *Die Umkehrfunktion* $\ln : ]0,\infty[ \to \mathbb{R}$ *zu* $\exp|_{\mathbb{R}} : \mathbb{R} \to ]0,\infty[$ *heißt* **natürlicher Logarithmus**.

**Lemma 13.8.4.**

*a)* $\ln : ]0,\infty[ \to \mathbb{R}$ *ist stetig.*

*b) Für alle* $x, y \in ]0,\infty[$ *gilt:*

$$\ln(xy) = \ln(x) + \ln(y) \tag{13.8.3}$$

*Beweis.*

a) Da $\exp : \mathbb{R} \to ]0,\infty[$ streng monoton steigend und stetig ist, gilt für $a, b \in \mathbb{R}$ mit $a < b$ nach dem Zwischenwertsatz:

$$(\exp|_{\mathbb{R}})(]a,b[) = ]e^a, e^b[.$$

Weil nach Lemma 13.5.10 jedes offene $\emptyset \neq U \subseteq \mathbb{R}$ abzählbare Vereinigung offener Intervalle ist, ist dann auch $(\exp|_{\mathbb{R}})(U)$ als Vereinigung offener Intervalle wieder offen. Weil ln die Umkehrfunktion von $\exp|_{\mathbb{R}}$ ist, gilt $\ln^{-1}(U) = (\exp|_{\mathbb{R}})(U)$ für alle (offenen) $U \subseteq \mathbb{R}$ und damit ist ln stetig.

b) Da ln die Umkehrfunktion von $\exp|_{\mathbb{R}}$ ist, gilt für alle $x, y \in ]0, \infty[$:

$$e^{\ln(x)+\ln(y)} = e^{\ln(x)} e^{\ln(y)} = xy = e^{\ln(xy)},$$

und da $\exp|_{\mathbb{R}}$ bijektiv ist, ist das zu $\ln(xy) = \ln(x) + \ln(y)$ äquivalent.

$\square$

**Definition 13.8.5.** *Für $a > 0$ und $z \in \mathbb{C}$ wird durch* $\begin{array}{rcl} a^z : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & a^z := e^{z \ln a} \end{array}$ *die Exponentialfunktion zur Basis $a > 0$ definiert.*

**Lemma 13.8.6.** *Die Funktion* $\begin{array}{rcl} \cos|_{[0,2]} : [0, 2] & \to & \mathbb{R} \\ x & \mapsto & \cos x \end{array}$ *ist streng monoton fallend und hat genau eine Nullstelle. Wir definieren $\frac{\pi}{2}$ als diese Nullstelle.*

*Beweis.* Für $x \in ]0, 2]$ sind die Folgen $\left(\frac{x^{2k}}{(2k)!}\right)_{k\in\mathbb{N}}$ und $\left(\frac{x^{2k+1}}{(2k+1)!}\right)_{k\in\mathbb{N}_0}$ in $[0, \infty[$ monoton fallend mit Grenzwert 0. Nach dem Leibnizkriterium für alternierende Reihen folgt damit $1 - \frac{x^2}{2} \le \cos x \le 1 - \frac{x^2}{2} + \frac{x^4}{24}$ und $x - \frac{x^3}{6} \le \sin x \le x$ für $x \in [0, 2]$. Daher ist $\sin x > 0$ für alle $x \in ]0, 2]$ und daher für $x, y \in [0, 2], x < y$ nach (6.6.19)

$$\cos x - \cos y = 2 \sin \frac{y + x}{2} \sin \frac{y - x}{2} > 0,$$

also $\cos|_{[0,2]} : [0, 2] \to \mathbb{R}$ streng monoton fallend, hat also höchstens eine Nullstelle in $[0, 2]$. Wegen $\cos 0 = 1$ und $-1 = 1 - \frac{2^2}{2} \le \cos 2 \le 1 - \frac{2^2}{2} + \frac{2^4}{24} = -\frac{1}{3}$ folgt mit dem Zwischenwertsatz, daß $\cos|_{[0,2]} : [0, 2] \to \mathbb{R}$ genau eine Nullstelle in $]0, 2[$ besitzt. $\square$

**Satz 13.8.7.**

- *Für die Vielfachen von $\frac{\pi}{2}$ bekommen wir folgende Tabelle:*

| $x$ | $0$ | $\frac{\pi}{2}$ | $\pi$ | $\frac{3\pi}{2}$ | $2\pi$ |
|---|---|---|---|---|---|
| $e^{ix}$ | $1$ | $i$ | $-1$ | $-i$ | $1$ |
| $\cos x$ | $1$ | $0$ | $-1$ | $0$ | $1$ |
| $\sin x$ | $0$ | $1$ | $0$ | $-1$ | $0$ |

- *Für alle $z \in \mathbb{C}$ gilt:*

$$\begin{array}{rclrclrcl} e^{z+i\frac{\pi}{2}} & = & ie^z & e^{z+i\pi} & = & -e^z & e^{z+i2\pi} & = & e^z \\ \cos(z + \frac{\pi}{2}) & = & -\sin z & \cos(z + \pi) & = & -\cos z & \cos(z + 2\pi) & = & \cos z \\ \sin(z + \frac{\pi}{2}) & = & \cos z & \sin(z + \pi) & = & -\sin z & \sin(z + 2\pi) & = & \sin z \end{array} \tag{13.8.4}$$

*Beweis.* Einsetzen von $x = 0$ in die Potenzreihen ergibt $\cos 0 = 1$ und $\sin 0 = 0$. Wegen $(\cos x)^2 + (\sin x)^2 = 1$ und $\cos(\frac{\pi}{2}) = 0$ folgt $(\sin(\frac{\pi}{2}))^2 = 1$. Mit $\frac{\pi}{2} \in ]0, 2[$ und $\sin x > 0$ für alle $x \in ]0, 2[$ folgt schließlich $\sin(\frac{\pi}{2}) = 1$. Aus der Eulerschen Formel folgt daher $e^{i\frac{\pi}{2}} =$

$\cos(\frac{\pi}{2}) + i\sin(\frac{\pi}{2}) = i$. Mit der Funktionalgleichung (6.6.1) für die Exponentialfunktion folgt: $e^{i\pi} = e^{i\frac{\pi}{2}+i\frac{\pi}{2}} = e^{i\frac{\pi}{2}} \cdot e^{i\frac{\pi}{2}} = i^2 = -1$ und ebenso $e^{i\frac{3\pi}{2}} = e^{i\pi} \cdot e^{i\frac{\pi}{2}} = -i$ und $e^{i2\pi} = e^{i\pi} \cdot e^{i\pi} = (-1)^2 = 1$. Über die Eulersche Formel ergeben sich daraus die Werte von $\sin x$ und $\cos x$ für $x = \pi, \frac{3\pi}{2}, 2\pi$. Aus der Funktionalgleichung (6.6.1) und der Tabelle folgt sofort $e^{z+i\frac{\pi}{2}} = ie^z$, $e^{z+i\pi} = -e^z$ und $e^{z+i2\pi} = e^z$ und wie zB. durch

$$\cos(z + \frac{\pi}{2}) = \frac{1}{2}\left(e^{i(z+\frac{\pi}{2})} + e^{-i(z+\frac{\pi}{2})}\right) = \frac{1}{2}(ie^{iz} - ie^{-iz}) = \sin(-z) = -\sin(z)$$

folgen auch die restlichen Gleichungen in (13.8.4). □

**Satz 13.8.8.** *Jede komplexe Zahl $z \in \mathbb{C}$ besitzt eine Polardarstellung*

$$z = |z|e^{i\varphi} \tag{13.8.5}$$

*mit $\varphi \in \mathbb{R}$. Ist $z \neq 0$, so ist das Argument $\varphi$ von $z$ bis auf ein ganzzahliges Vielfaches von $2\pi$ eindeutig bestimmt.*

*Beweis.* Nach (13.8.4) ist $\cos(\frac{\pi}{2} - x) = \sin x = -\cos(\frac{\pi}{2} + x)$, also $\cos : [0, \pi] \to [-1, 1]$ streng monoton fallend und daher injektiv. Nach dem Zwischenwertsatz ist diese Funktion auch surjektiv und daher gibt es die Umkehrabbildung $\arccos : [-1, 1] \to [0, \pi]$. Für $z \neq 0$ mit $\operatorname{Im} z \geq 0$ betrachte $w := \frac{z}{|z|} = a + ib$, dann ist $a^2 + b^2 = 1$ und $b \geq 0$. Setze $\varphi := \arccos a$, so ist $\varphi \in [0, \pi]$ und $\cos\varphi = a$, $\sin\varphi \geq 0$. Wegen $a^2 + b^2 = 1$ und $b \geq 0$ folgt $b = \sin\varphi$ und das zeigt $w = a + ib = \cos\varphi + i\sin\varphi = e^{i\varphi}$ oder $z = |z|e^{i\varphi}$. Im Fall $\operatorname{Im} z < 0$ hat $\overline{z}$ nach dem eben behandelten Fall eine Polardarstellung $\overline{z} = |z|e^{i\varphi}$ mit $\varphi \in [0, \pi]$ und daher ist $z = |z|e^{-i\varphi}$ die gesuchte Polardarstellung. Sind nun $z = |z|e^{i\varphi} = |z|e^{i\theta}$ zwei Polardarstellungen von $z$, so ist $e^{i(\varphi-\theta)} = 1$ oder $\varphi - \theta = 2k\pi$ mit $k \in \mathbb{Z}$. □

**Bemerkung 13.8.9.** Sind $w = |w|e^{i\theta} \neq 0$ und $z = |z|e^{i\varphi} \neq 0$, dann ist $wz = |w|\cdot|z|e^{i(\theta+\varphi)}$, dh. in Polarkoordinaten ist die Multiplikation von komplexen Zahlen die Multiplikation der Beträge und die Addition der Argumente. Als Anwendung kann man für jedes $w \in \mathbb{C}\backslash\{0\}$ und $n \in \mathbb{N}$ alle $n$ Lösungen von

$$z^n = w \tag{13.8.6}$$

folgendermaßen bestimmen:

- Schreibe $w = |w|e^{i\theta}$ in Polarkoordinaten, dann ist $z_0 := \sqrt[n]{|w|}e^{i\frac{\theta}{n}}$ eine Lösung von (13.8.6).
- Wegen $2\pi i$- Periodizität der Exponentialfunktion sind für alle $k \in \mathbb{Z}$ die komplexen Zahlen $z_k := \sqrt[n]{|w|}e^{i\frac{\theta+2\pi k}{n}}$ weitere Lösungen von (13.8.6).
- Nach dem Fundamentalsatz der Algebra besitzt (13.8.6) (mit Vielfachheiten gezählt) genau $n$ verschiedene Lösungen. Die Menge $\{z_k : k \in \mathbb{Z}\}$ von Lösungen von (13.8.6) besitzt jetzt genau $n$ Elemente, nämlich $z_0, z_1, ..., z_{n-1}$,

## 13.9 Stetige lineare Abbildungen

**Satz 13.9.1.** *Es seien $V$ und $W$ normierte Räume und $T : V \to W$ sei $\mathbb{K}$-linear. Dann sind äquivalent:*

*a) $T : V \to W$ ist gleichmäßig stetig.*

*b) $T : V \to W$ ist stetig.*

*c) $T : V \to W$ ist stetig im Punkt $\mathbf{0}$.*

*d) Zu jedem $\varepsilon > 0$ existiert $\delta = \delta(\varepsilon) > 0$ mit $\|T(x)\| \le \varepsilon$ für alle $x \in V$ mit $\|x\| < \delta$.*

*e) Es gibt ein $C \in ]0, \infty[$ mit $\|T(x)\| \le C\|x\|$ für alle $x \in V$.*

*Beweis.* Offenbar gelten die Implikationen *a)* $\Rightarrow$ *b)* $\Rightarrow$ *c)*.

c)$\Rightarrow$d) $T(\mathbf{0}) = \mathbf{0}$, daher steht in (d) die Stetigkeit von $T$ in $\mathbf{0}$.

d)$\Rightarrow$e) Es sei $\delta := \delta(1)$ aus d) gewählt, dann ist $\left\| \frac{x\delta}{\|x\|} \right\| \le \delta$ für alle $x \in V \setminus \{\mathbf{0}\}$, also folgt aus der Voraussetzung in d) die Abschätzung $\left\| T\left(\frac{x\delta}{\|x\|}\right) \right\| \le 1$ oder $\|T(x)\| \le \frac{1}{\delta}\|x\|$ für $x \neq \mathbf{0}$. Für $x = \mathbf{0}$ ist diese Abschätzung klar, denn $T(\mathbf{0}) = \mathbf{0}$.

e)$\Rightarrow$a) Sind $x, y \in V$, so gilt wegen der Linearität von $T$ und der Voraussetzung in e):

$$\|T(x) - T(y)\| = \|T(x - y)\| \le C\|x - y\|,$$

dh. $T$ ist gleichmäßig stetig – man kann $\delta(\varepsilon) = \frac{1}{C}\varepsilon$ in der Definition 13.2.3 wählen.

□

**Definition 13.9.2.** *Sind $V$ und $W$ normierte Räume, dann heißt*

$$L(V, W) := \{T : V \to W : T \textit{linear; es gibt } C \in ]0, \infty[ \textit{ mit } \|T(x)\| \le C\|x\| \textit{ für alle } x \in V\} \tag{13.9.1}$$

*der Raum der* **stetigen linearen** *Abbildungen von $V$ nach $W$.*

**Satz 13.9.3.** *Es seien $V, W$ normierte Räume,*

- *dann definiert*

$$|||T||| := \inf\{C \in ]0, \infty[ : \|T(x)\| \le C\|x\| \textit{ für alle } x \in V\} \tag{13.9.2}$$

$$= \sup\left\{\frac{\|T(x)\|}{\|x\|} : x \in V \setminus \{\mathbf{0}\}\right\} = \sup\{\|T(x)\| : \|x\| = 1\} \tag{13.9.3}$$

$$= \sup\{\|T(x)\| : \|x\| \le 1\} \tag{13.9.4}$$

*eine Norm auf $L(V, W)$, die sogenannte* **Operatornorm**.

- *Für jedes $x \in V$ ist*

$$\|T(x)\| \leq |||T||| \cdot \|x\|. \tag{13.9.5}$$

- *Ist $W$ ein Banachraum, so ist $(L(V,W), |||\cdot|||)$ ein Banachraum.*
- *Ist $V = W$, so ist*

$$|||TS||| \leq |||T||| \cdot |||S||| \tag{13.9.6}$$

*für alle $S, T \in L(V,V)$.*

*Beweis.* Für $T \neq 0 \in L(V,W)$ ist laut Definition

$$\emptyset \neq \{C \in ]0,\infty[ : \|T(x)\| \leq C\|x\| \text{ für alle } x \in V\} =: M,$$

also existiert $|||T||| := \inf(M) \in [0,\infty[$ als Infimum der nichtleeren nach unten beschränkten Teilmenge $M$ von $\mathbb{R}$. Für jedes $C \in M$ ist $\dfrac{\|T(x)\|}{\|x\|} \leq C$ für jedes $x \in V\backslash\{\mathbf{0}\}$, also ist auch $s := \sup\{\frac{\|T(x)\|}{\|x\|} : x \in V\backslash\{\mathbf{0}\}\} \leq C$ und damit nach Infimumsbildung $s \leq |||T|||$. Andererseits ist $s \in M$ und damit $s = |||T|||$, was die erste Gleichheit in (13.9.3) zeigt. Für alle $x \neq \mathbf{0}$ ist $\frac{\|T(x)\|}{\|x\|} = \left\|T(\frac{x}{\|x\|})\right\|$ und daher folgt aus $\left\|\dfrac{x}{\|x\|}\right\| = 1$ die zweite Gleichheit in (13.9.3). Für jedes $\mathbf{0} \neq x \in V$ mit $\|x\| \leq 1$ ist $\|T(x)\| = \|x\| \cdot \left\|T(\frac{x}{\|x\|})\right\| \leq \left\|T(\frac{x}{\|x\|})\right\|$, also gilt auch (13.9.4). Aus (13.9.2) folgt $\|T(x)\| \leq |||T||| \cdot \|x\|$ durch Bilden des Infimums über $C$. Für jedes $\lambda \in \mathbb{K}$, $x \in V$ und $T, S \in L(V,W)$ ist $\|(\lambda T)(x)\| = |\lambda| \cdot \|T(x)\|$ und $\|(T+S)(x)\| = \|T(x) + S(x)\| \leq \|T(x)\| + \|S(x)\|$, also gilt nach Bilden von $\sup_{\|x\|=1}$ auch $|||\lambda T||| = |\lambda| \cdot |||T|||$ und $|||T+S||| \leq |||T||| + |||S|||$.

Ist $W$ ein Banachraum und $(T_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $L(V,W)$, so ist für jedes $x \in V$ auch $(T_n(x))_{n\in\mathbb{N}}$ eine Cauchyfolge in $W$, deshalb existiert $T(x) := \lim\limits_{n\to\infty} T_n(x) \in W$. Definiere damit $\begin{array}{rcl} T : V & \to & W \\ x & \mapsto & T(x) = \lim\limits_{n\to\infty} T_n(x) \end{array}$, dann gilt:

- Für $\lambda, \mu \in \mathbb{K}$ und $x, y \in V$ gilt wegen Linearität der Abbildungen $T_n$:

$$T(\lambda x + \mu y) = \lim_{n\to\infty} T_n(\lambda x + \mu y) = \lim_{n\to\infty} (\lambda T_n(x) + \mu T_n(y))$$

  und da Addition und Skalarmultiplikation stetig sind folgt daraus

$$\lim_{n\to\infty} (\lambda T_n(x) + \mu T_n(y)) = \lambda \lim_{n\to\infty} T_n(x) + \mu \lim_{n\to\infty} T_n(y) = \lambda T(x) + \mu T(y)$$

  dh. $T$ ist linear.

- Zu jedem $\varepsilon > 0$ wähle $N = N(\varepsilon) \in \mathbb{N}$ mit $|||T_n - T_m||| < \frac{\varepsilon}{2}$ für alle $m, n \geq N(\varepsilon)$. Für $x \in V$ mit $\|x\| \leq 1$ wähle $M = M(x,\varepsilon) \in \mathbb{N}$ mit $\|T_m(x) - T(x)\| < \frac{\varepsilon}{2}$ für alle $m \geq M(x,\varepsilon)$. Daher ist

$$\|T_n(x) - T(x)\| \leq \|T_n(x) - T_m(x)\| + \|T_m(x) - T(x)\| \leq \|T_n(x) - T_m(x)\| + \frac{\varepsilon}{2} < \varepsilon$$

falls $m \geq \max\{M(x,\varepsilon), N(\varepsilon)\}$ und $n \geq N(\varepsilon)$ und da man immer $m \geq \max\{M(x,\varepsilon), N(\varepsilon)\}$ wählen kann, gilt

$$\|T_n(x) - T(x)\| \leq \varepsilon \tag{13.9.7}$$

für alle $n \geq N(\varepsilon)$. Nach Bilden von $\sup_{\|x\|=1}$ in (13.9.7) folgt, daß $T_n - T$ und damit $T$ stetig ist und $|||T_n - T||| \leq \varepsilon$ und damit $|||T||| \leq |||T_n||| + |||T_n - T|||$, und $T_n \overset{n\to\infty}{\longrightarrow} T$ bezüglich $|||\cdot|||$.

Ist $V = W$ und sind $S, T \in L(V,V)$, dann ist

$$\|(TS)(x)\| \leq |||T||| \cdot \|S(x)\| \leq |||T||| \cdot |||S||| \cdot \|x\|,$$

also folgt $|||TS||| \leq |||T||| \cdot |||S|||$ nach Bilden von $\sup_{\|x\|=1}$. □

**Satz 13.9.4.** *Es sei $m, n \in \mathbb{N}$, dann sind alle $\mathbb{K}$-linearen Abbildungen $F : \mathbb{K}^n \to \mathbb{K}^m$ stetig bzgl. den Normtopologien.*

*Beweis.* Es sei $\underline{e}_1, ..., \underline{e}_n$ die Standardbasis von $\mathbb{K}^n$, $\underline{e}_1, ..., \underline{e}_m$ die Standardbasis von $\mathbb{K}^m$ und

$$A = \begin{pmatrix} \langle \underline{e}_1, F(\underline{e}_1)\rangle & \cdots & \langle \underline{e}_1, F(\underline{e}_n)\rangle \\ \vdots & & \vdots \\ \langle \underline{e}_m, F(\underline{e}_1)\rangle & \cdots & \langle \underline{e}_m, F(\underline{e}_n)\rangle \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

die darstellende Matrix von $F$ bezüglich der Standardbasen. Für $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$ gilt:

$$\begin{aligned} \|F(\underline{x})\|_\infty &= \left\| \sum_{k=1}^n x_k F(\underline{e}_k) \right\|_\infty = \left\| \sum_{k=1}^n x_k \sum_{l=1}^m \langle \underline{e}_l, F(\underline{e}_k)\rangle \underline{e}_l \right\|_\infty \\ &= \max \left\{ \left| \sum_{k=1}^n x_k a_{lk} \right| : l = 1, ..., m \right\} \leq \max\{\sum_{k=1}^n |a_{lk}| : l = 1, ..., m\} \|\underline{x}\|_\infty \end{aligned}$$

dh. $F$ ist stetig bezüglich der $\|\cdot\|_\infty$-Normen und da alle Normen auf $\mathbb{K}^n$ bzw. $\mathbb{K}^m$ äquivalent sind, ist $F$ auch bezüglich aller anderen Normen stetig. □

**Satz 13.9.5.** *Es sei $(X, \|\cdot\|)$ ein normierter Raum und $T \in L(X,X)$, so gilt:*

*a) Konvergiert $\sum_{n=0}^\infty T^n$ in $L(X,X)$, dann ist $\mathrm{id}_X - T$ invertierbar und*

$$(\mathrm{id}_X - T)^{-1} = \sum_{n=0}^\infty T^n \in L(X,X). \tag{13.9.8}$$

*b) Ist $X$ ein Banachraum und $|||T||| < 1$, dann ist*

$$(\mathrm{id}_X - T)^{-1} = \sum_{n=0}^{\infty} T^n$$

*und*

$$|||(\mathrm{id}_X - T)^{-1}||| \leq \frac{1}{1 - |||T|||}.$$

*Beweis.* Es sei $S_m := \sum_{n=0}^{m} T^n$ , dann ist $(\mathrm{id}_X - T)S_m = S_m(\mathrm{id}_X - T) = \mathrm{id}_X - T^{m+1}$. Da die Reihe $\sum_{n=0}^{\infty} T^n$ konvergiert, bildet $(T^n)_{n\in\mathbb{N}}$ eine Nullfolge in $L(X, X)$ und dann ist

$$\mathrm{id}_X = \lim_{m\to\infty} (\mathrm{id}_X - T^{m+1}) = \lim_{m\to\infty} (\mathrm{id}_X - T)S_m = (\mathrm{id}_X - T) \lim_{m\to\infty} S_m$$

und genauso folgt

$$\mathrm{id}_X = \lim_{m\to\infty} S_m(\mathrm{id}_X - T),$$

also gilt

$$(\mathrm{id}_X - T)^{-1} = \sum_{n=0}^{\infty} T^n.$$

Ist $X$ ein Banachraum und $|||T||| < 1$, so gilt $\sum_{n=0}^{\infty} |||T^n||| \leq \sum_{n=0}^{\infty} |||T|||^n < \infty$, also ist $\sum_{n=0}^{\infty} T^n$ als absolut konvergente Reihe in einem Banachraum auch konvergent und aus

$$|||\sum_{n=0}^{\infty} T^n||| \leq \sum_{n=0}^{\infty} |||T|||^n = \frac{1}{1 - |||T|||}$$

folgt die Normabschätzung. □

# Kapitel 14

# Integration

## 14.1 $\sigma$-Algebren und meßbare Abbildungen

**Definition 14.1.1.** *Es sei $\emptyset \neq X$ eine Menge. Dann heißt $\mathcal{A} \subseteq \mathcal{P}(X)$ eine* **$\sigma$-Algebra** *auf $X$, wenn*

*a)* $\emptyset, X \in \mathcal{A}$.

*b) Für jedes $A \in \mathcal{A}$ ist auch $X \backslash A \in \mathcal{A}$.*

*c) Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$, dann ist auch $\bigcup\limits_{n\in\mathbb{N}} A_n \in \mathcal{A}$.*

*$(X, \mathcal{A})$ heißt ein* **Meßraum**, *wenn $\mathcal{A}$ eine $\sigma$-Algebra auf $X$ ist.*

**Bemerkung 14.1.2.** Es sei $(X, \mathcal{A})$ ein Meßraum und $(V_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$, dann ist

- $\bigcap\limits_{n\in\mathbb{N}} (X \backslash V_n) = X \backslash \left( \bigcup\limits_{n\in\mathbb{N}} V_n \right) \in \mathcal{A}$.
- $\bigcap\limits_{n\in\mathbb{N}} V_n = \bigcap\limits_{n\in\mathbb{N}} (X \backslash (X \backslash V_n)) = X \backslash \left( \bigcup\limits_{n\in\mathbb{N}} (X \backslash V_n) \right) \in \mathcal{A}$.

**Beispiel 14.1.3.** Es sei $X \neq \emptyset$ eine Menge, dann sind

1. $\mathcal{A}_1 = \{\emptyset, X\}$
2. $\mathcal{A}_2 = \mathcal{P}(X)$
3. $\mathcal{A}_3 = \{U \subseteq X : U \text{ abzählbar oder } X \backslash U \text{ abzählbar }\}$

$\sigma$-Algebren auf $X$.

*Beweis.* Für $\mathcal{A}_1$ und $\mathcal{A}_2$ ist alles klar. Ist $X$ abzählbar, so gilt $\mathcal{A}_3 = \mathcal{P}(X)$ und nur wenn $X$ nicht abzählbar ist, dann ist für $\mathcal{A}_3$ ist noch etwas zu zeigen:

- $\emptyset, X \in \mathcal{A}$, da $\emptyset = X \backslash X$ abzählbar ist.

- Ist $U \in \mathcal{A}$ und $U$ abzählbar, so ist $X \backslash U \in \mathcal{A}$, da $U = X \backslash (X \backslash U)$ abzählbar ist.
- Ist $U \in \mathcal{A}$ und $X \backslash U$ abzählbar, so ist $X \backslash U \in \mathcal{A}$.
- Ist $(U_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{A}$. Sind alle $U_n$ abzählbar, so ist auch $\bigcup\limits_{n \in \mathbb{N}} U_n$ abzählbar nach Korollar 2.2.8. Gibt es ein $m \in \mathbb{N}$, so daß $X \backslash U_m$ abzählbar ist, dann ist auch $X \backslash \bigcup\limits_{n \in \mathbb{N}} U_n = \bigcap\limits_{n \in \mathbb{N}} (X \backslash U_n) \subseteq X \backslash U_m$ abzählbar, also ist auch in diesem Fall $\bigcup\limits_{n \in \mathbb{N}} U_n \in \mathcal{A}$.

$\square$

**Beispiel 14.1.4.** Es sei $\mathcal{A}$ eine $\sigma$-Algebra auf $X$.

a) Ist $Y \subseteq X$, so ist

$$\mathcal{A}_Y := \{Y \cap V : V \in \mathcal{A}\} \subseteq \mathcal{P}(Y)$$

eine $\sigma$-Algebra auf $Y$.

b) Ist $T : Z \to X$ eine Abbildung, so ist $\{T^{-1}(V) : V \in \mathcal{A}\} \subseteq \mathcal{P}(Z)$ eine $\sigma$-Algebra auf $Z$.

c) Ist $f : X \to Y$ eine Abbildung, so ist

$$\mathcal{C} := \{V \subseteq Y : f^{-1}(V) \in \mathcal{A}\} \tag{14.1.1}$$

eine $\sigma$-Algebra auf $Y$.

*Beweis.*

a) $\emptyset = Y \cap \emptyset \in \mathcal{A}_Y$ und $Y = Y \cap X \in \mathcal{A}_Y$. Ist $U = Y \cap V \in \mathcal{A}_Y$, dann ist auch $Y \backslash U = Y \cap (X \backslash V) \in \mathcal{A}_Y$. Ist $(U_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{A}_Y$ mit $U_n = Y \cap V_n$, dann ist auch $\bigcup\limits_{n \in \mathbb{N}} U_n = Y \cap \left( \bigcup\limits_{n \in \mathbb{N}} V_n \right) \in \mathcal{A}_Y$ und daher ist $\mathcal{A}_Y$ eine $\sigma$-Algebra.

b) Aus den Rechenregeln für das Urbild, nämlich $T^{-1}(\emptyset) = \emptyset, T^{-1}(X) = Z, T^{-1}(X \backslash V) = Z \backslash T^{-1}(V)$ und $T^{-1} \left( \bigcup\limits_{n \in \mathbb{N}} V_n \right) = \bigcup\limits_{n \in \mathbb{N}} T^{-1}(V_n)$ folgt, daß $\{T^{-1}(V) : V \in \mathcal{A}\}$ eine $\sigma$-Algebra ist.

c) $f^{-1}(\emptyset) = \emptyset \in \mathcal{A}$ und $f^{-1}(Y) = X \in \mathcal{A}$, daher sind $\emptyset, Y \in \mathcal{C}$. Für $V \in \mathcal{C}$ ist $f^{-1}(Y \backslash V) = X \backslash f^{-1}(V) \in \mathcal{A}$ (als Komplement von $f^{-1}(V)$), also ist $Y \backslash V \in \mathcal{C}$. Ist $(V_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{C}$, dann ist $f^{-1} \left( \bigcup\limits_{n \in \mathbb{N}} V_n \right) = \bigcup\limits_{n \in \mathbb{N}} f^{-1}(V_n) \in \mathcal{A}$, also $\bigcup\limits_{n \in \mathbb{N}} V_n \in \mathcal{C}$.

$\square$

**Lemma 14.1.5.** *Ist $X$ eine Menge, $\mathcal{A}_i, i \in I$ eine Familie von $\sigma$-Algebren auf $X$, so ist*

$$\bigcap_{i \in I} \mathcal{A}_i = \{V \subseteq X : V \in \mathcal{A}_i \text{ für alle } i \in I\}$$

*eine $\sigma$-Algebra auf $X$.*

*Beweis.* Da jedes $\mathcal{A}_i$ für $i \in I$ eine $\sigma$-Algebra auf $X$ ist, so gilt $\emptyset, X \in \mathcal{A}_i$ für jedes $i \in I$, also $\emptyset, X \in \bigcap_{i\in I} \mathcal{A}_i$. Ist $V \in \bigcap_{i\in I} \mathcal{A}_i$, so folgt $V \in \mathcal{A}_i$ für jedes $i \in I$ daher $X\backslash V \in \mathcal{A}_i$ für jedes $i \in I$, also $X\backslash V \in \bigcap_{i\in I} \mathcal{A}_i$. Ist $(V_n)_{n\in\mathbb{N}}$ eine Folge in $\bigcap_{i\in I} \mathcal{A}_i$, so ist für jedes $i \in I$ auch $(V_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}_i$ und damit $\bigcup_{n\in\mathbb{N}} V_n \in \mathcal{A}_i$ für jedes $i \in I$ und daher $\bigcup_{n\in\mathbb{N}} V_n \in \bigcap_{i\in I} \mathcal{A}_i$. □

**Definition 14.1.6.** *Ist $X$ eine Menge, $\mathcal{E} \subseteq \mathcal{P}(X)$, so heißt*

$$\sigma(\mathcal{E}) := \bigcap_{\substack{\mathcal{E}\subseteq\mathcal{A}\subseteq\mathcal{P}(X)\\ \mathcal{A}\sigma-Algebra}} \mathcal{A}$$

*die von $\mathcal{E}$* **erzeugte $\sigma$-Algebra**. *Ist $(X, \mathcal{O})$ ein topologischer Raum, dann heißt*

$$\mathcal{B} = \mathcal{B}(X) := \sigma(\mathcal{O})$$

*die* **Borelsche $\sigma$-Algebra** *auf $(X, \mathcal{O})$.*

**Bemerkung 14.1.7.** Es sei $\emptyset \neq X$ eine Menge und $\mathcal{E}, \mathcal{F} \subseteq \mathcal{P}(X)$. Nach Definition der erzeugten $\sigma$-Algebra ist $\sigma(\mathcal{E})$ die kleinste $\sigma$-Algebra auf $X$, die das Mengensystem $\mathcal{E}$ enthält. Es ist immer $\{\emptyset, X\} \subseteq \sigma(\mathcal{E}) \subseteq \mathcal{P}(X)$ und nach Einsetzen in die Definition folgt

$$\mathcal{E} \subseteq \mathcal{F} \Rightarrow \sigma(\mathcal{E}) \subseteq \sigma(\mathcal{F}). \tag{14.1.2}$$

Ferner ist für jede $\sigma$-Algebra $\mathcal{A}$ auf $X$ nach Definition $\sigma(\mathcal{A}) = \mathcal{A}$, also insbesondere

$$\sigma(\mathcal{E}) = \sigma(\sigma(\mathcal{E})). \tag{14.1.3}$$

Wird nichts anderes gesagt, so faßt man einen topologischen Raum $(X, \mathcal{O}_X)$ versehen mit der Borelschen $\sigma$-Algebra $\mathcal{B}(X) = \sigma(\mathcal{O}_X)$ als Meßraum $(X, \mathcal{B}(X))$ auf.

**Lemma 14.1.8.** *Es sei $(X, \mathcal{O})$ ein topologischer Raum und $\mathcal{A} := \{X\backslash V : V \in \mathcal{O}\}$ die abgeschlossenen Teilmengen von $X$ und $\mathcal{K} := \{K \subseteq X : K \text{ kompakt }\}$ die Menge der kompakten Teilmengen von $X$, dann gilt:*

*a) $\mathcal{B}(X) = \sigma(\mathcal{A})$.*

*b) Es sei $\mathcal{E}$ eine abzählbare Basis der Topologie $\mathcal{O}$, dann gilt $\mathcal{B}(X) = \sigma(\mathcal{E})$.*

*c) Ist $X$ hausdorffsch und gibt es eine Folge $(K_n)_{n\in\mathbb{N}}$ von kompakten Mengen mit $X = \bigcup_{n\in\mathbb{N}} K_n$, dann gilt $\mathcal{B}(X) = \sigma(\mathcal{K})$.*

*Beweis.*

a) Da eine $\sigma$-Algebra mit einer Menge auch deren Komplement enthält, folgt $\mathcal{A} \subseteq \sigma(\mathcal{O})$ und $\mathcal{O} \subseteq \sigma(\mathcal{A})$, daher ist $\mathcal{B}(X) = \sigma(\mathcal{O}) \subseteq \sigma(\sigma(\mathcal{A})) = \sigma(\mathcal{A}) \subseteq \sigma(\sigma(\mathcal{O})) = \sigma(\mathcal{O})$.

b) Da $\mathcal{E}$ nur abzählbar viele Teilmengen von $X$ enthält, sind nach Definition 14.1.1 auch alle offenen Mengen – als Vereinigung von abzählbar vielen Elementen aus $\mathcal{E}$ – in $\sigma(\mathcal{E})$ enthalten, damit ist $\sigma(\mathcal{E}) \supseteq \mathcal{B}(X)$, die andere Inklusion folgt aus $\mathcal{E} \subseteq \mathcal{O}$.

c) Nach Lemma 13.6.2 ist jede kompakte Teilmenge eines hausdorffschen Raums abgeschlossen, daher ist $\sigma(\mathcal{K}) \subseteq \sigma(\mathcal{A}) = \mathcal{B}(X)$. Da jede abgeschlossene Teilmenge eines kompakten Raums nach Lemma 13.6.2 wieder kompakt ist, schreibt sich wegen $X = \bigcup_{n\in\mathbb{N}} K_n$ jede abgeschlossene Menge $A = \bigcup_{n\in\mathbb{N}} (A \cap K_n)$ als abzählbare Vereinigung von kompakten Mengen, also $\mathcal{A} \subseteq \sigma(\mathcal{K})$. $\square$

**Definition 14.1.9.** *Es seien $(X, \mathcal{A})$ und $(Y, \mathcal{B})$ Meßräume, dann heißt eine Abbildung $f : X \to Y$ $\mathcal{A} - \mathcal{B}-$**meßbar**, wenn $f^{-1}(B) \in \mathcal{A}$ für jedes $B \in \mathcal{B}$ erfüllt ist.*

**Beispiel 14.1.10.** Ist $(X, \mathcal{A})$ ein Meßraum, $A \subseteq X$, dann ist $A \in \mathcal{A}$, genau dann wenn

$$\begin{array}{rcl} \mathbf{1}_A : X & \to & \mathbb{R} \\ x & \mapsto & \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{falls } x \in X \setminus A \end{cases} \end{array}$$

meßbar ist.

*Beweis.* Für $B \in \mathcal{B}(\mathbb{R})$ ist

$$\mathbf{1}_A^{-1}(B) = \begin{cases} \emptyset & \text{falls } 0 \notin B \text{ und } 1 \notin B \\ A & \text{falls } 0 \notin B \text{ und } 1 \in B \\ X \setminus A & \text{falls } 0 \in B \text{ und } 1 \notin B \\ X & \text{falls } 0 \in B \text{ und } 1 \in B \end{cases}$$

Da $\mathcal{A}$ eine $\sigma$-Algebra auf $X$ ist, gilt immer $\emptyset, X \in \mathcal{A}$ und $X \setminus A \in \mathcal{A}$ genau dann wenn $A \in \mathcal{A}$ ist. $\square$

**Satz 14.1.11.** *Es seien $(X, \mathcal{A})$ und $(Y, \mathcal{B})$ Meßräume und $\mathcal{E} \subseteq \mathcal{P}(Y)$ mit $\sigma(\mathcal{E}) = \mathcal{B}$. Dann sind äquivalent:*

*a) $f : X \to Y$ ist $\mathcal{A} - \mathcal{B}-$meßbar*

*b) Für jedes $B \in \mathcal{E}$ ist $f^{-1}(B) \in \mathcal{A}$*

*Beweis.*

a) $\Rightarrow$ b) folgt aus der Definition der Meßbarkeit, da $\mathcal{E} \subseteq \mathcal{B}$.

b) $\Rightarrow$ a) $\mathcal{C} := \{V \subseteq Y : f^{-1}(V) \in \mathcal{A}\}$ ist nach Beispiel 14.1.4 eine $\sigma$-Algebra auf $Y$ mit $\mathcal{E} \subseteq \mathcal{C}$. Daher ist $\mathcal{B} = \sigma(\mathcal{E}) \subseteq \mathcal{C}$ und daher $f$ meßbar. $\square$

**Korollar 14.1.12.** *Es seien $(X, \mathcal{O}_X)$ und $(Y, \mathcal{O}_Y)$ topologische Räume und $f : X \to Y$ stetig; dann ist $f$ $\mathcal{B}(X) - \mathcal{B}(Y)-$meßbar.*

*Beweis.* Nach Satz 13.1.15 ist $f^{-1}(V) \in \mathcal{O}_X \subseteq \mathcal{B}(X)$ für jedes $V \in \mathcal{O}_Y$ und wegen $\sigma(\mathcal{O}_Y) = \mathcal{B}(Y)$ ist $f$ nach Satz 14.1.11 meßbar. $\square$

**Lemma 14.1.13.**

$$\begin{array}{rcl} \mathcal{I} & := & \{]a,b] : a,b \in \mathbb{R}, a \le b\} \\ \mathcal{I}_{\mathbb{Q}} & := & \{]a,b] : a,b \in \mathbb{Q}, a \le b\} \\ \mathcal{I}^\circ & := & \{]a,b[ : a,b \in \mathbb{R}, a \le b\} \\ \mathcal{I}^\circ_{\mathbb{Q}} & := & \{]a,b[ : a,b \in \mathbb{Q}, a \le b\} \\ \overline{\mathcal{I}} & := & \{[a,b] : a,b \in \mathbb{R}, a \le b\} \cup \{\emptyset\} \\ \overline{\mathcal{I}}_{\mathbb{Q}} & := & \{[a,b] : a,b \in \mathbb{Q}, a \le b\} \cup \{\emptyset\} \end{array}$$

*sind* **durchschnittsstabile Erzeugendensysteme** *von* $\mathcal{B}(\mathbb{R})$ *(enthalten also mit zwei Mengen auch deren Durchschnitt).*

*Beweis.* Die Mengensysteme enthalten mit zwei Elementen auch den Durchschnitt, sind also durchschnittsstabil. Nach Lemma 13.5.10 ist jede offene Teilmenge $U$ von $\mathbb{R}$ eine abzählbare Vereinigung von paarweise disjunkten offenen Intervallen $U = \bigcup_{n\in\mathbb{N}} ]a_n, b_n[$ mit $a_n < b_n$, wobei $a_n = -\infty$ und $b_m = \infty$ jeweils einmal erlaubt sind. Wegen $]a_m, \infty[= \bigcup_{k\in\mathbb{Z}} (]a_m, \infty[\cap]k-1, k+1[)$ ist auch $]a_m, \infty[$ (und analog $]-\infty, b_n[$), somit auch $U$ eine abzählbare Vereinigung von offenen Intervallen aus $\mathcal{I}^\circ$. Folglich ist $\mathcal{O} \subseteq \sigma(\mathcal{I}^\circ)$, also folgt $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{O}) \subseteq \sigma(\mathcal{I}^\circ) \subseteq \sigma(\mathcal{O})$ aus $\mathcal{I}^\circ \subseteq \mathcal{O}$. Nach Lemma 4.1.3 gibt es für zwei reelle Zahlen $a, b \in \mathbb{R}$ mit $a < b$ ein $q \in \mathbb{Q}\cap]a, b[$, daher ist – zB. mit der Wahl von $p \in \mathbb{Q}\cap]a, a + \frac{1}{n}[$ und $q \in \mathbb{Q}\cap]b - \frac{1}{n}, b[$ – dann

$$]a, b[= \bigcup_{\substack{p,q\in\mathbb{Q} \\ a<p\leq q<b}} ]p, q[ \tag{14.1.4}$$

eine abzählbare Vereinigung von offenen Intervallen mit rationalen Endpunkten, also $]a, b[\in \sigma(\mathcal{I}^\circ_\mathbb{Q})$ und damit $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{I}^\circ_\mathbb{Q})$. Da Intervalle $[a, b]$ abgeschlossen sind, folgt

$$\sigma(\overline{\mathcal{I}}_\mathbb{Q}) \subseteq \sigma(\overline{\mathcal{I}}) \subseteq \sigma(\mathcal{A}) = \sigma(\mathcal{O}) = \mathcal{B}(\mathbb{R}).$$

Jedes halboffene Intervall $]a, b]$ hat eine Darstellung als $]a, b] =]a, c[\cap[d, b]$ mit $c > b$ und $d < a$ und ist deswegen in $\sigma(\mathcal{O}) = \mathcal{B}(\mathbb{R})$ enthalten, damit ist

$$\sigma(\mathcal{I}_\mathbb{Q}) \subseteq \sigma(\mathcal{I}) \subseteq \mathcal{B}(\mathbb{R}).$$

Analog zu (14.1.4) gilt

$$]a, b[= \bigcup_{\substack{p,q\in\mathbb{Q} \\ a<p\leq q<b}} ]p, q] = \bigcup_{\substack{p,q\in\mathbb{Q} \\ a<p\leq q<b}} [p, q]$$

also folgt $\mathcal{I}^\circ_\mathbb{Q} \subseteq \mathcal{I}^\circ \subseteq \sigma(\mathcal{I}_\mathbb{Q}) \subseteq \sigma(\mathcal{I})$ und $\mathcal{I}^\circ_\mathbb{Q} \subseteq \mathcal{I}^\circ \subseteq \sigma(\overline{\mathcal{I}}_\mathbb{Q}) \subseteq \sigma(\overline{\mathcal{I}})$, was den Beweis beendet. □

**Lemma 14.1.14.**

$$\mathcal{B}(\widehat{\mathbb{R}}) = \{B \subseteq \widehat{\mathbb{R}} : \mathbb{R} \cap B \in \mathcal{B}(\mathbb{R})\} = \{B \cup E : B \in \mathcal{B}(\mathbb{R}), E \subseteq \{-\infty, \infty\}\} \tag{14.1.5}$$

*ist die kleinste* $\sigma-$*Algebra auf* $\widehat{\mathbb{R}}$*, die* $\{-\infty\}$, $\{\infty\}$ *und alle offenen Teilmengen von* $\mathbb{R}$ *enthält.*

$$\mathcal{B}([0, \infty]) = \{B \subseteq [0, \infty] : B \cap [0, \infty[\in \mathcal{B}([0, \infty[) \tag{14.1.6}$$

*Beweis.* Offenbar gilt

$$\{B \subseteq \widehat{\mathbb{R}} : \mathbb{R} \cap B \in \mathcal{B}(\mathbb{R})\} = \{B \cup E : B \in \mathcal{B}(\mathbb{R}), E \subseteq \{-\infty, \infty\}\} =: \mathcal{B}$$

und weil

- $\emptyset = \emptyset \cup \emptyset \in \mathcal{B}$, $\widehat{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\} \in \mathcal{B}$
- mit $A = B \cup E \in \mathcal{B}$ – also $B \in \mathcal{B}(\mathbb{R})$ und $E \subseteq \{-\infty, \infty\}$ ist

$$\widehat{\mathbb{R}} \backslash A = (\mathbb{R} \backslash B) \cup (\{-\infty, \infty\} \backslash E) \in \mathcal{B}.$$

- Ist $(A_n)_{n\in\mathbb{N}} = (B_n \cup E_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{B}$ – also $B_n \in \mathcal{B}(\mathbb{R})$ und $E_n \subseteq \{-\infty, \infty\}$, dann ist

$$\bigcup_{n\in\mathbb{N}} A_n = \left(\bigcup_{n\in\mathbb{N}} B_n\right) \cup \left(\bigcup_{n\in\mathbb{N}} E_n\right) \in \mathcal{B}$$

ist $\mathcal{B}$ eine $\sigma-$Algebra auf $\widehat{\mathbb{R}}$. Wegen $\sigma_{\mathbb{R}}(\mathcal{O}_{\mathbb{R}}) = \mathcal{B}(\mathbb{R})$ ist $\mathcal{B}$ die kleinste $\sigma-$Algebra auf $\widehat{\mathbb{R}}$, die $\{-\infty\}, \{\infty\}$ und alle offenen Teilmengen von $\mathbb{R}$ enthält. Da nach Lemma 5.6.5 eine reelle Folge $(x_n)_{n\in\mathbb{N}}$ genau dann gegen einen Grenzwert $x \in \mathbb{R}$ bzgl. $|\cdot|$ konvergiert, wenn sie bzgl. der Metrik $d$ auf $\widehat{\mathbb{R}}$ aus Lemma 5.6.5 konvergiert, gilt

$$\mathcal{O}_{\widehat{\mathbb{R}}} \cap \mathbb{R} = \{U \cap \mathbb{R} : U \in \mathcal{O}_{\widehat{\mathbb{R}}}\} = \mathcal{O}_{\mathbb{R}}$$

für die von $|\cdot|$ auf $\mathbb{R}$ erzeugte Topologie $\mathcal{O}_{\mathbb{R}}$ bzw. die auf $\widehat{\mathbb{R}}$ durch $d$ erzeugte Topologie $\mathcal{O}_{\widehat{\mathbb{R}}}$. Ferner ist nach Lemma 13.5.10 jedes $U \in \mathcal{O}_{\mathbb{R}}$ eine abzählbare Vereinigung paarweise disjunkter offener Intervalle, wobei je einmal $]-\infty, b[$ oder $]a, \infty[$ vorkommen kann; insbesondere ist für $V \in \mathcal{O}_{\widehat{\mathbb{R}}}$ dann auch $V \cap \mathbb{R}$ von dieser Form. Ist $\infty \in V$, so ist $\infty$ kein isolierter Punkt von $V$ – denn sonst ist $\{\infty\}$ als einelementige Menge eine abgeschlossene Zusammenhangskomponente von $V$. Da jede Umgebung von $\infty$ bzgl. $\mathcal{O}_{\widehat{\mathbb{R}}}$ ein Intervall $]a, \infty]$ enthält, gibt es in einer Zerlegung $V \cap \mathbb{R} = \bigcup\limits_{n\in N} I_n$ von $V \cap \mathbb{R}$ in abzählbar viele paarweise disjunkte offene Intervalle eines der Form $]a_n, \infty[$. Analog läßt sich bei $-\infty$ argumentieren und deshalb gibt es zu $V \in \mathcal{O}_{\widehat{\mathbb{R}}}$ eine Folge $I_n, n \in N$ von Intervallen

- $I_n = ]a_n, b_n[$, $a_n, b_n \in \mathbb{R}$, $a_n < b_n$
- und $m \in N$ mit $I_m = ]-\infty, b_m[$ oder $I_m = [-\infty, b_m[$, falls $V$ in $\mathbb{R}$ nicht nach unten beschränkt ist
- und $k \in N$ mit $I_k = ]a_k, \infty[$ oder $I_k = ]a_k, \infty]$, falls $V$ in $\mathbb{R}$ nicht nach oben beschränkt ist.

mit $V = \bigcup\limits_{n\in\mathbb{N}} V_n$. Dies zeigt schon

$$\mathcal{O}_{\widehat{\mathbb{R}}} \subseteq \{B \cup E : B \in \mathcal{B}(\mathbb{R}), E \subseteq \{-\infty, \infty\}\} = \mathcal{B}$$

und da $\mathcal{B}$ eine $\sigma-$Algebra auf $\widehat{\mathbb{R}}$ ist, gilt $\mathcal{B}(\widehat{\mathbb{R}}) \subseteq \mathcal{B}$. Da im metrischen Raum $\widehat{\mathbb{R}}$ alle einelementigen Mengen abgeschlossen sind, sind insbesondere $\{-\infty\}, \{\infty\}$ abgeschlossene Teilmengen von $\widehat{\mathbb{R}}$ und daher $\mathbb{R}$ eine offene Teilmenge von $\widehat{\mathbb{R}}$. Es gilt also $\mathcal{O}_{\mathbb{R}} \subseteq \mathcal{O}_{\widehat{\mathbb{R}}}$. Weil $\mathcal{B}$ die kleinste $\sigma-$Algebra auf $\widehat{\mathbb{R}}$ ist, die $\{-\infty\}, \{\infty\} \in \mathcal{B}(\widehat{\mathbb{R}})$ und alle Mengen aus $\mathcal{O}_{\mathbb{R}}$ enthält, folgt daraus auch $\mathcal{B} \subseteq \mathcal{B}(\widehat{\mathbb{R}})$. (14.1.6) zeigt man analog. □

**Lemma 14.1.15.**

$$\begin{aligned}\mathcal{J} &:= \{]\alpha,\infty] : \alpha \in \mathbb{R}\}\\ \mathcal{J}_{\mathbb{Q}} &:= \{]\alpha,\infty] : \alpha \in \mathbb{Q}\}\end{aligned}$$

*sind Erzeugendensysteme für $\mathcal{B}(\widehat{\mathbb{R}})$.*

*Beweis.* Nach Lemma 14.1.14 ist $\mathcal{B}(\widehat{\mathbb{R}}) = \{B \subseteq \widehat{\mathbb{R}} : B \cap \mathbb{R} \in \mathcal{B}(\mathbb{R})\}$ und damit gilt $]\alpha,\infty] \in \mathcal{B}(\widehat{\mathbb{R}})$ für jedes $\alpha \in \mathbb{R}$. Damit folgt

$$\sigma_{\widehat{\mathbb{R}}}(\mathcal{J}_{\mathbb{Q}}) \subseteq \sigma_{\widehat{\mathbb{R}}}(\mathcal{J}) \subseteq \sigma_{\widehat{\mathbb{R}}}(\mathcal{B}(\widehat{\mathbb{R}})) = \mathcal{B}(\widehat{\mathbb{R}}).$$

Wegen $]\alpha,\beta] = ]\alpha,\infty] \setminus ]\beta,\infty]$ ist $\mathcal{I}_{\mathbb{Q}} \subseteq \sigma_{\widehat{\mathbb{R}}}(\mathcal{J}_{\mathbb{Q}})$ und $\mathcal{I} \subseteq \sigma_{\widehat{\mathbb{R}}}(\mathcal{J})$ also

$$\sigma_{\mathbb{R}}(\mathcal{I}_{\mathbb{Q}}) = \sigma_{\mathbb{R}}(\mathcal{I}) = \mathcal{B}(\mathbb{R}) \subseteq \sigma_{\widehat{\mathbb{R}}}(\mathcal{J}_{\mathbb{Q}}) \subseteq \sigma_{\widehat{\mathbb{R}}}(\mathcal{J}).$$

Da $\mathcal{B}(\widehat{\mathbb{R}})$ nach Lemma 14.1.14 die kleinste $\sigma$-Algebra auf $\widehat{\mathbb{R}}$ ist, die $\mathcal{B}(\mathbb{R})$ enthält, ist auch $\mathcal{B}(\widehat{\mathbb{R}}) = \sigma_{\widehat{\mathbb{R}}}(\mathcal{B}(\mathbb{R})) \subseteq \sigma_{\widehat{\mathbb{R}}}(\mathcal{J}_{\mathbb{Q}}) \subseteq \sigma_{\widehat{\mathbb{R}}}(\mathcal{J})$. □

**Korollar 14.1.16.** *Ist $(X, \mathcal{A})$ ein Meßraum, dann sind äquivalent:*

*a)* $f : X \to \widehat{\mathbb{R}}$ *ist meßbar.*

*b) Für jedes $\alpha \in \mathbb{R}$ ist $f^{-1}(]\alpha,\infty]) \in \mathcal{A}$.*

*Beweis.* Nach Lemma 14.1.15 ist $\{]\alpha,\infty] : \alpha \in \mathbb{R}\}$ ein Erzeugendensystem von $\mathcal{B}(\widehat{\mathbb{R}})$, daher folgt die Äquivalenz aus Satz 14.1.11. □

**Korollar 14.1.17.** *Es sei $(X, \mathcal{A})$ ein Meßraum und $(f_n : X \to \widehat{\mathbb{R}})_{n\in\mathbb{N}}$ eine Folge von meßbaren Funktionen, dann sind auch*

$$\begin{aligned}\sup_{n\in\mathbb{N}} f_n : X &\to \widehat{\mathbb{R}} \\ x &\mapsto \sup_{n\in\mathbb{N}} f_n(x)\end{aligned} \qquad \begin{aligned}\inf_{n\in\mathbb{N}} f_n : X &\to \widehat{\mathbb{R}} \\ x &\mapsto \inf_{n\in\mathbb{N}} f_n(x)\end{aligned}$$

$$\begin{aligned}\limsup_{n\in\mathbb{N}} f_n : X &\to \widehat{\mathbb{R}} \\ x &\mapsto \limsup_{n\in\mathbb{N}} f_n(x)\end{aligned} \qquad \begin{aligned}\liminf_{n\in\mathbb{N}} f_n : X &\to \widehat{\mathbb{R}} \\ x &\mapsto \liminf_{n\in\mathbb{N}} f_n(x)\end{aligned}$$

*meßbar. Gilt die punktweise Konvergenz $f = \lim_{n\to\infty} f_n$, so ist $f$ meßbar.*

*Beweis.* Für jedes $\alpha \in \mathbb{R}$ ist

$$\{x \in X : \sup_{n\in\mathbb{N}} f_n(x) > \alpha\} = \bigcup_{n\in\mathbb{N}} \{x \in X : f_n(x) > \alpha\} \in \mathcal{A},$$

da nach Voraussetzung $\{x \in X : f_n(x) > \alpha\} = f_n^{-1}(]\alpha,\infty]) \in \mathcal{A}$. Da nach Lemma 14.1.15 das System $\{]\alpha,\infty] : \alpha \in \mathbb{R}\}$ ein Erzeugendensystem von $\mathcal{B}(\widehat{\mathbb{R}})$ ist, folgt daraus die Meßbarkeit von $\sup_{n\in\mathbb{N}} f_n$ nach Satz 14.1.11. Wegen $\inf_{n\in\mathbb{N}} f_n = -\sup_{n\in\mathbb{N}}(-f_n)$, $\limsup_{n\to\infty} f_n = \inf_{n\in\mathbb{N}} \sup_{k\geq n} f_k$ und $\liminf_{n\to\infty} f_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} f_k$ sind auch $\inf_{n\in\mathbb{N}} f_n, \limsup_{n\to\infty} f_n$ und $\liminf_{n\to\infty} f_n$ meßbar. Ist $f = \lim_{n\to\infty} f_n$, so gilt $f = \liminf_{n\to\infty} f_n = \limsup_{n\to\infty} f_n$ und daher ist $f$ meßbar. □

**Bemerkung 14.1.18.** Für jedes $x \in \mathbb{R}$ heißt $x_+ := \max\{x, 0\}$ der **Positivteil** von $x$ und $x_- := \max\{-x, 0\}$ der **Negativteil** von $x$. Offenbar ist

$$\begin{aligned} x &= x_+ - x_- && (14.1.7)\\ |x| &= x_+ + x_- && (14.1.8)\\ x_- &= (-x)_+ && (14.1.9)\\ x_+ \cdot x_- &= 0 && (14.1.10) \end{aligned}$$

für jedes $x \in \mathbb{R}$.

**Korollar 14.1.19.** *Ist $(X, \mathcal{A})$ ein Meßraum und $f : X \to \mathbb{R}$ meßbar, dann sind der Positivteil* $\begin{array}{rcl} f_+ : X & \to & [0, \infty[ \\ x & \mapsto & f_+(x) := \max\{0, f(x)\} \end{array}$ *von $f$ und auch der Negativteil* $\begin{array}{rcl} f_- : X & \to & [0, \infty[ \\ x & \mapsto & f_-(x) := \max\{0, -f(x)\} \end{array}$ *von $f$ meßbar.*

**Definition 14.1.20.** *Es sei $(X, \mathcal{A})$ ein Meßraum. Eine meßbare Funktion $f : X \to \mathbb{C}$ heißt ($\mathcal{A}$-) Stufenfunktion, wenn $f(X)$ eine endliche Menge ist.*

**Bemerkung 14.1.21.** Es sei $(X, \mathcal{A})$ ein Meßraum, dann ist $f : X \to \mathbb{C}$ genau dann eine $\mathcal{A}$-Stufenfunktion, wenn es $n \in \mathbb{N}$, paarweise verschiedene $\alpha_1, ..., \alpha_n \in \mathbb{C}$ und paarweise disjunkte Mengen $A_1, ..., A_n \in \mathcal{A}$ gibt mit $f = \sum_{k=1}^{n} \alpha_k \mathbf{1}_{A_k}$.

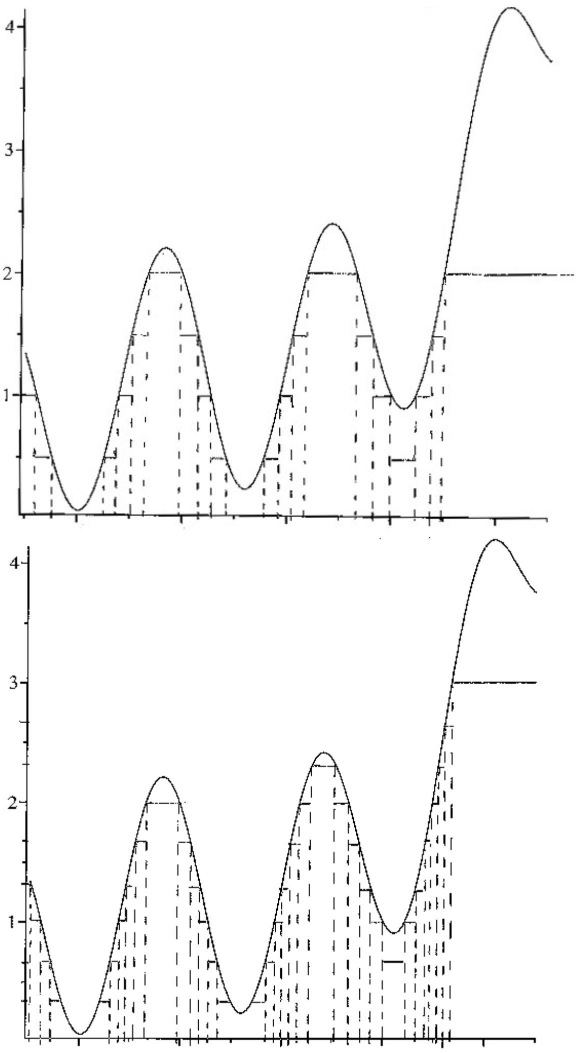
**Satz 14.1.22.** *Es sei $(X, \mathcal{A})$ ein Meßraum und $f : X \to [0, \infty]$ meßbar. Dann gibt es eine Folge $(f_n : X \to [0, \infty[)_{n\in\mathbb{N}}$ von $\mathcal{A}$-Stufenfunktionen mit $f(x) = \lim_{n\to\infty} f_n(x)$ für jedes $x \in X$ und $0 \le f_n(x) \le f_{n+1}(x) \le f(x)$ für alle $n \in \mathbb{N}$ und $x \in X$.*

*Beweis.*

$$\begin{array}{rcl} g_n : X & \to & \mathbb{R} \\ x & \mapsto & \begin{cases} \frac{m-1}{n} & \text{falls } \frac{m-1}{n} \le f(x) < \frac{m}{n} \text{ mit } m \in \mathbb{N}, 1 \le m \le n^2 \\ n & \text{falls } f(x) \ge n \end{cases} \end{array} \qquad (14.1.11)$$

ist offenbar eine $\mathcal{A}$-Stufenfunktion mit $g_n \le f$, $f(x) - g_n(x) \le \frac{1}{n}$ für alle $x \in \{x \in X : f(x) \le n\}$. Nach Korollar 14.1.17 ist auch $f_n := \sup_{k\le n} g_k$ meßbar, also eine $\mathcal{A}$-Stufenfunktion. Es gilt $0 \le g_n(x) \le f_n(x) \le f_{n+1}(x) \le f(x)$ für alle $x \in X$ und $n \in \mathbb{N}$ und daraus folgt die punktweise Konvergenz $\lim_{n\to\infty} f_n = f$. □

Die Bilder zeigen zu einer Funktion die drei ersten Näherungen durch die Stufenfunktionen $g_1$, $g_2$ und $g_3$, wie sie im Beweis von Satz 14.1.22 angegeben werden. Beim Vergleich des zweiten und dritten Bildes sieht man auch rein optisch, das $(g_n)_{n\in\mathbb{N}}$ keine monotone Folge von Funktionen ist.


4
3
2
1
4
3
2
1

## 14.2 Maße

**Definition 14.2.1.** *Es sei $(X, \mathcal{A})$ ein Meßraum. Eine Abbildung $\mu : \mathcal{A} \to [0, \infty]$ heißt ein* **Maß**, *wenn*

*a)* $\mu(\emptyset) = 0$

*b) Für jede Folge von paarweise disjunkten Elementen $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ gilt:*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n). \qquad \sigma - \textbf{Additivität} \tag{14.2.1}$$

*In diesem Fall heißt $(X, \mathcal{A}, \mu)$ ein* **Maßraum**. *Ist $(X, \mathcal{A}, \mu)$ ein Maßraum, dann heißt*

- $\mu$ *ein* **Wahrscheinlichkeitsmaß** *und $(X, \mathcal{A}, \mu)$ ein* **Wahrscheinlichkeitsraum**, *wenn $\mu(X) = 1$ ist.*
- $\mu$ *ein* **endliches Maß**, *wenn $\mu(X) < \infty$ ist.*
- $\mu$ *ein* **$\sigma$-endliches Maß**, *wenn es eine Folge $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ gibt mit $A_1 \subseteq A_2 \subseteq \ldots$, $X = \bigcup\limits_{n\in\mathbb{N}} A_n$ und $\mu(A_k) < \infty$ für alle $k \in \mathbb{N}$.*

*Ist $(X, \mathcal{O})$ ein topologischer Raum und $\mathcal{B}(X) = \sigma(\mathcal{O})$ die Borelsche $\sigma$-Algebra auf $X$, dann heißt jedes Maß $\mu : \mathcal{B}(X) \to [0, \infty]$ mit $\mu(K) < \infty$ für jede kompakte Teilmenge $K \subseteq X$ ein* **Borelmaß** *auf $X$.*

**Beispiel 14.2.2.**

a) Es sei $\emptyset \neq X$ eine Menge, dann definiert

$$\begin{aligned} \nu : \mathcal{P}(X) &\to [0, \infty] \\ Y &\mapsto \begin{cases} |Y| := \text{ Zahl der Elemente von } Y & \text{falls } Y \text{ endlich} \\ \infty & \text{falls } Y \text{ nicht endlich} \end{cases} \end{aligned}$$

das **Zählmaß** auf $X$.

b) Ist $\emptyset \neq X$ eine endliche Menge mit $n := |X|$, dann definiert

$$\begin{aligned} \nu_n : \mathcal{P}(X) &\to [0, 1] \\ Y &\mapsto \tfrac{|Y|}{n} \end{aligned} \tag{14.2.2}$$

ein Wahrscheinlichkeitsmaß auf $\mathcal{P}(X)$. Für die Modellierung des Münzwurfs ist $X = \{Kopf, Zahl\}$ und schon durch $\nu_2(\{Kopf\}) = \nu_2(\{Zahl\}) = \frac{1}{2}$ das Wahrscheinlichkeitsmaß $\nu_2$ auf $\mathcal{P}(X)$ eindeutig festgelegt. Ebenso ist beim fairen Würfel $X = \{1, 2, 3, 4, 5, 6\}$ und $\nu_6$ das zugehörige Wahrscheinlichkeitsmaß.

c) Ist $X$ eine abzählbar unendliche Menge, dann ist das Zählmaß $\nu : \mathcal{P}(X) \to [0, \infty]$ $\sigma$-endlich. Denn nach Definition einer abzählbar unendlichen Menge gibt es eine bijektive Abbildung $\varphi : X \to \mathbb{N}$, daher bildet $A_n := \varphi^{-1}(\{1, \ldots, n\})$ eine Folge in $\mathcal{P}(X)$ mit $n = |A_n| = \nu(A_n)$, $A_1 \subseteq A_2 \subseteq \ldots$ und $X = \bigcup\limits_{n\in\mathbb{N}} A_n$.

d) Es sei $(X, \mathcal{A})$ ein Meßraum und $x \in X$. Dann definiert

$$\begin{aligned} \delta_x : \mathcal{A} &\to \{0,1\} \\ A &\mapsto \delta_x(A) := \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{falls } x \in X \backslash A \end{cases} \end{aligned} \tag{14.2.3}$$

ein Maß auf $\mathcal{A}$, das **Diracmaß**, denn $\delta_x(\emptyset) = 0$ und ist $(A_n)_{n\in\mathbb{N}}$ eine Folge von paarweise disjunkten Mengen in $\mathcal{A}$, so ist $x$ in höchstens einer dieser Mengen enthalten, also

$$\delta_x\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \delta_x(A_n) = \begin{cases} 1 & \text{falls es } k \in \mathbb{N} \text{ mit } x \in A_k \text{ gibt} \\ 0 & \text{sonst} \end{cases}$$

e) Sind $(X, \mathcal{A})$ und $(Y, \mathcal{B})$ Meßräume, $f : X \to Y$ $\mathcal{A} - \mathcal{B}-$meßbar und $\mu : \mathcal{A} \to [0, \infty]$ ein Maß. Dann definiert

$$\begin{aligned} \mu \circ f^{-1} : \mathcal{B} &\to [0, \infty] \\ B &\mapsto \mu(f^{-1}(B)) \end{aligned} \tag{14.2.4}$$

ein Maß, das sogenannte **Bildmaß** von $\mu$ bei $f$. Denn $(\mu \circ f^{-1})(\emptyset) = \mu(\emptyset) = 0$ und für jede Folge $(B_n)_{n\in\mathbb{N}}$ von paarweise disjunkten Elementen von $\mathcal{B}$ ist $(f^{-1}(B_n))_{n\in\mathbb{N}} = (\{x \in X : f(x) \in B_n\})_{n\in\mathbb{N}}$ eine Folge von paarweise disjunkten Mengen in $\mathcal{A}$ und weil $\mu$ ein Maß ist folgt

$$(\mu \circ f^{-1})\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} f^{-1}(B_n)\right) = \sum_{n\in\mathbb{N}} \mu(f^{-1}(B_n)) = \sum_{n\in\mathbb{N}} (\mu \circ f^{-1})(B_n)$$

**Bemerkung 14.2.3.** Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum, $A, B \in \mathcal{A}$ und $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$, dann gilt:

- $\mu$ ist **additiv**, dh. für paarweise disjunkte Mengen $A_1, ..., A_n \in \mathcal{A}$ gilt

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k).$$

- $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$
- Ist $A \subseteq B$, dann ist $\mu(A) \leq \mu(B)$.
- Ist $A \subseteq B$ und $\mu(B) < \infty$, dann ist $\mu(B \backslash A) = \mu(B) - \mu(A)$.
- $\mu$ ist **$\sigma$-subadditiv:** $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n)$

*Beweis.* Durch Betrachten der Folge $(A_1, ..., A_n, \emptyset, \emptyset, ....)$ aus paarweise disjunkten Mengen folgt die Additivität von $\mu$ aus der $\sigma$-Additivität. $A = (A \cap B) \cup (A \cap (X \backslash B))$,

$B = (A \cap B) \cup (B \cap (X \backslash A))$ und $A \cup B = (A \cap B) \cup (A \cap (X \backslash B)) \cup (X \backslash A) \cap B)$ sind Zerlegungen von $A, B$ und $A \cup B$ in disjunkte Teilmengen. Daher gilt wegen $(\sigma-)$Additivität von $\mu$:

$$\begin{aligned} \mu(A \cup B) + \mu(A \cap B) &= \mu(A \cap B) + \mu(A \cap (X \backslash B)) + \mu((X \backslash A) \cap B) + \mu(A \cap B) \\ &= (\mu(A \cap B) + \mu(A \cap (X \backslash B))) + (\mu(A \cap B) + \mu((X \backslash A) \cap B)) \\ &= \mu(A) + \mu(B). \end{aligned} \tag{14.2.5}$$

Sind $A, B \in \mathcal{A}$ mit $A \subseteq B$, dann ist $B = (B \cap A) \cup (B \cap (X \backslash A)) = A \cup (B \backslash A)$ eine Zerlegung von $B$ in paarweise disjunkte Mengen, also

$$\mu(B) = \mu(A) + \mu(B \backslash A) \geq \mu(A). \tag{14.2.6}$$

Ist dann $\mu(B) < \infty$, dann ist auch $\mu(A), \mu(B \backslash A) < \infty$ also folgt $\mu(B \backslash A) = \mu(B) - \mu(A)$ aus (14.2.6). Zur Folge $(A_n)_{n \in \mathbb{N}}$ definiere induktiv die Folge $(B_n)_{n \in \mathbb{N}}$ durch

$$\begin{aligned} B_1 &:= A_1 \\ &\vdots \\ B_{n+1} &:= A_{n+1} \cap \left( X \backslash \bigcup_{k=1}^{n} B_k \right) \end{aligned} \tag{14.2.7}$$

dann ist $(B_n)_{n \in \mathbb{N}}$ eine Folge von paarweise disjunkten Mengen in $\mathcal{A}$ mit $\bigcup\limits_{n \in \mathbb{N}} A_n = \bigcup\limits_{n \in \mathbb{N}} B_n$ und $B_n \subseteq A_n$, also folgt aus der $\sigma$-Additivität von $\mu$:

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \mu\left(\bigcup_{n \in \mathbb{N}} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

□

**Definition 14.2.4.** *Es sei $(X, \mathcal{A})$ ein Meßraum und $\mu : \mathcal{A} \to [0, \infty]$ mit $\mu(\emptyset) = 0$, dann heißt $\mu$*

- *$\sigma$-stetig von unten, wenn für jede Folge $(B_n)_{n \in \mathbb{N}}$ in $\mathcal{A}$ mit $B_1 \subseteq B_2 \subseteq \ldots$ und $B := \bigcup\limits_{n \in \mathbb{N}} B_n$ gilt (in $[0, \infty]$):*

$$\lim_{n \to \infty} \mu(B_n) = \mu(B). \tag{14.2.8}$$

- *$\sigma$-stetig von oben, wenn für jede Folge $(C_n)_{n \in \mathbb{N}}$ in $\mathcal{A}$ mit $C_1 \supseteq C_2 \supseteq \ldots$ und $C := \bigcap\limits_{n \in \mathbb{N}} C_n$ gilt:*

$$\lim_{n \to \infty} \mu(C_n) = \mu(C). \tag{14.2.9}$$

- *$\emptyset$-stetig, wenn für jede Folge $(D_n)_{n \in \mathbb{N}}$ in $\mathcal{A}$ mit $D_1 \supseteq D_2 \supseteq \ldots$ und $\emptyset = \bigcap\limits_{n \in \mathbb{N}} D_n$ gilt:*

$$\lim_{n \to \infty} \mu(D_n) = 0. \tag{14.2.10}$$

**Lemma 14.2.5.** *Es sei $(X, \mathcal{A})$ ein Meßraum und $\mu : \mathcal{A} \to [0, \infty]$ mit $\mu(\emptyset) = 0$. Dann sind äquivalent:*

*a) $\mu$ ist ein Maß.*

*b) $\mu$ ist additiv und $\sigma$-stetig von unten.*

*Ist $\mu(X) < \infty$, dann ist das äquivalent zu*

*c) $\mu$ ist additiv und $\sigma$-stetig von oben.*

*d) $\mu$ ist additiv und $\emptyset$-stetig.*

*Beweis.*

a)⇒b) Ist $\mu$ $\sigma$-additiv und $(B_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$ mit $B_1 \subseteq B_2 \subseteq ...$, dann ist die induktiv gebildete Folge

$$\begin{aligned} \tilde{B}_1 &:= B_1 \\ \vdots & \quad \vdots \\ \tilde{B}_{n+1} &:= B_{n+1} \setminus \left( \bigcup_{k=1}^{n} \tilde{B}_k \right) \end{aligned}$$

eine Folge $(\tilde{B}_n)_{n\in\mathbb{N}}$ von paarweise disjunkten Mengen in $\mathcal{A}$ mit $B_n = \bigcup\limits_{k=1}^{n} \tilde{B}_k$. Aus der $\sigma$-Additivität von $\mu$ folgt:

$$\mu(B_n) = \mu\left( \bigcup_{k=1}^{n} \tilde{B}_k \right) = \sum_{k=1}^{n} \mu(\tilde{B}_k) \overset{n\to\infty}{\longrightarrow} \sum_{k=1}^{\infty} \mu(\tilde{B}_k) = \mu(B).$$

b)⇒a) Ist $\mu$ additiv und $\sigma$-stetig von unten und $(A_n)_{n\in\mathbb{N}}$ eine Folge von paarweise disjunkten Mengen in $\mathcal{A}$, dann ist durch $B_n := A_1 \cup ... \cup A_n$ eine Folge in $\mathcal{A}$ mit $B_1 \subseteq B_2 \subseteq ...$ und $\bigcup\limits_{n\in\mathbb{N}} B_n = \bigcup\limits_{n\in\mathbb{N}} A_n$ gegeben. Wegen $\sigma$-Stetigkeit von unten ist

$$\mu\left( \bigcup_{n\in\mathbb{N}} A_n \right) = \mu\left( \bigcup_{n\in\mathbb{N}} B_n \right) = \lim_{n\to\infty} \mu\left( \bigcup_{k=1}^{n} B_k \right) = \lim_{n\to\infty} \mu\left( \bigcup_{k=1}^{n} A_k \right)$$

und aus der Additivität von $\mu$ folgt

$$\mu\left( \bigcup_{k=1}^{n} A_k \right) = \sum_{k=1}^{n} \mu(A_k),$$

was zusammengenommen

$$\mu\left( \bigcup_{n\in\mathbb{N}} A_n \right) \overset{n\to\infty}{\longleftarrow} \mu\left( \bigcup_{k=1}^{n} A_k \right) = \sum_{k=1}^{n} \mu(A_k) \overset{n\to\infty}{\longrightarrow} \sum_{k=1}^{\infty} \mu(A_k)$$

die $\sigma$-Additivität von $\mu$ zeigt.

Es sei nun $\mu(X) < \infty$ und $\mu$ additiv. Ist $\mu$ $\sigma$-stetig von unten und $(C_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$ mit $C_1 \supseteq C_2 \supseteq \ldots$ und $C := \bigcap_{n\in\mathbb{N}} C_n$, dann gelten für $B_n := X\backslash C_n$ die Inklusionen $B_1 \subseteq B_2 \subseteq \ldots$ und $B := \bigcup_{n\in\mathbb{N}} B_n = X\backslash \bigcap_{n\in\mathbb{N}} C_n = X\backslash C$ und aus der $\sigma$-Stetigkeit von unten folgt: $\lim_{n\to\infty} \mu(B_n) = \mu(B) = \mu(X\backslash C) = \mu(X) - \mu(C)$ (wobei die letzte Gleichheit wegen $\mu(X) < \infty$ nach Bemerkung 14.2.3 folgt). Da $C_n = X\backslash B_n$ ist, gilt $\lim_{n\to\infty} \mu(C_n) = \mu(X) - \lim_{n\to\infty} \mu(B_n) = \mu(C)$, dh. $\mu$ ist $\sigma$-stetig von oben. Aus der $\sigma$-Stetigkeit von oben folgt die $\emptyset$-Stetigkeit als Spezialfall $C = \emptyset$. Ist $\mu$ $\emptyset$-stetig und $(B_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$ mit $B_1 \subseteq B_2 \subseteq \ldots$ und $B := \bigcup_{n\in\mathbb{N}} B_n$, dann definiert $D_n := B\backslash B_n$ eine Folge in $\mathcal{A}$ mit $D_1 \supseteq D_2 \supseteq \ldots$ und $\bigcap_{n\in\mathbb{N}} D_n = \emptyset$ und daher folgt aus der $\emptyset$-Stetigkeit: $0 \xleftarrow{n\to\infty} \mu(D_n) = \mu(B\backslash B_n) = \mu(B) - \mu(B_n)$ also $\lim_{n\to\infty} \mu(B_n) = \mu(B)$. □

**Bemerkung 14.2.6.** Nimmt man sich das Wahrscheinlichkeitsmaß $\nu_2$ für den Münzwurf zum Vorbild, wo man durch die beiden Festlegungen von $\nu_2(\{Kopf\}) = \nu_2(\{Zahl\}) = \frac{1}{2}$ schon ein Wahrscheinlichkeitsmaß auf $\mathcal{P}(X)$ definieren konnte, so würde man sich wünschen, daß man sich durch $\mu([a,b]) := b - a$ ein Maß auf $\mathcal{P}(\mathbb{R})$ definieren könnte. Daß so eine Konstruktion unweigerlich zu Widersprüchen führt, wenn man so ein Maß auf ganz $\mathcal{P}(\mathbb{R})$ definieren will, das ist ein Spezialfall eines Satzes von Vitali (1905) und ein guter Grund, wieso man $\sigma$-Algebren $\mathcal{A} \subsetneq \mathcal{P}(X)$ betrachtet.

**Definition 14.2.7.** *Es sei $\emptyset \neq X$ eine Menge, $\mathcal{R} \subseteq \mathcal{P}(X)$ heißt ein* **Mengenring***, wenn*

- $\emptyset \in \mathcal{R}$
- *Für jedes $A, B \in \mathcal{R}$, $A \supseteq B$ ist $A\backslash B \in \mathcal{R}$.*
- *Zu $A, B \in \mathcal{R}$ ist $A \cup B \in \mathcal{R}$.*

*Gilt zusätzlich noch $X \in \mathcal{R}$, so heißt $\mathcal{R}$ eine* **Mengenalgebra***.*
*Eine Abbildung $\mu : \mathcal{R} \to [0, \infty]$ mit*

- $\mu(\emptyset) = 0$
- *Für je endlich viele paarweise disjunkte Mengen $A_1, \ldots, A_N \in \mathcal{R}$ gilt*

$$\mu\left(\bigcup_{k=1}^{N} A_k\right) = \sum_{k=1}^{N} \mu(A_k) \tag{14.2.11}$$

*heißt ein* **Inhalt***. Eine Abbildung $\mu : \mathcal{R} \to [0, \infty]$ mit*

- $\mu(\emptyset) = 0$
- *Für jede Folge von paarweise disjunkten Mengen $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{R}$, die noch zusätzlich $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{R}$ erfüllt, gilt*

$$\mu\left(\bigcup_{k\in\mathbb{N}} A_k\right) = \sum_{k\in\mathbb{N}} \mu(A_k) \tag{14.2.12}$$

*heißt ein* **Prämaß**. *Eine Funktion* $\mu^* : \mathcal{P}(X) \to [0, \infty]$ *mit*

- $\mu^*(\emptyset) = 0$
- *Für alle* $A \subseteq B \subseteq X$ *gilt:* $\mu^*(A) \leq \mu^*(B)$.
- *Für jede Folge* $(A_n)_{n\in\mathbb{N}}$ *in* $\mathcal{P}(X)$ *gilt*

$$\mu^* \left( \bigcup_{n\in\mathbb{N}} A_n \right) \leq \sum_{n\in\mathbb{N}} \mu^*(A_n) \qquad \sigma - \textbf{Subadditivität} \tag{14.2.13}$$

*heißt ein* **äußeres Maß**.

**Satz 14.2.8.** *Es sei* $(X, \mathcal{A})$ *ein Meßraum,* $\mathcal{E} \subseteq \mathcal{P}(X)$ *ein durchschnittsstabiles Erzeugendensystem von* $\mathcal{A}$ *und* $(E_n)_{n\in\mathbb{N}}$ *eine Folge in* $\mathcal{E}$ *mit* $\bigcup\limits_{n\in\mathbb{N}} E_n = X$. *Dann sind je zwei Maße* $\mu_1 : \mathcal{A} \to [0, \infty]$ *und* $\mu_2 : \mathcal{A} \to [0, \infty]$ *mit*

- $\mu_1(A) = \mu_2(A)$ *für alle* $A \in \mathcal{E}$.
- $\mu_1(E_n) = \mu_2(E_n) < \infty$ *für alle* $n \in \mathbb{N}$

*identisch auf ganz* $\mathcal{A}$.

**Satz 14.2.9** (Caratheodory)**.** *Es sei* $\emptyset \neq X$ *und* $\mu^*$ *ein äußeres Maß auf* $X$, *dann ist das Mengensystem*

$$\begin{aligned} \mathcal{A} &:= \{A \subseteq X : \mu^*(Q) \geq \mu^*(Q \cap A) + \mu^*(Q \cap (X \backslash A)) \text{ für alle } Q \subseteq X\} \qquad (14.2.14) \\ &= \{A \subseteq X : \mu^*(Q) \geq \mu^*(Q \cap A) + \mu^*(Q \cap (X \backslash A)) \text{ für alle } Q \subseteq X \text{ mit } \mu^*(Q) < \infty\} \end{aligned}$$

*eine* $\sigma$*-Algebra auf* $X$ *und die Restriktion* $\mu$ *von* $\mu^*$ *auf* $\mathcal{A}$, *also* $\begin{array}{rcl} \mu : \mathcal{A} & \to & [0, \infty] \\ A & \mapsto & \mu^*(A) \end{array}$ *ein Maß.*

**Satz 14.2.10.** *Es sei* $\emptyset \neq X$ *eine Menge und* $\mathcal{R} \subseteq \mathcal{P}(X)$ *ein Mengenring. Jedes Prämaß* $\nu$ *auf* $\mathcal{R}$ *läßt sich auf mindestens eine Weise zu einem äußeren Maß* $\mu^*$ *auf* $\mathcal{P}(X)$ *fortsetzen, so daß* $\mu := \mu^*|_{\sigma(\mathcal{R})}$ *ein Maß auf der von* $\mathcal{R}$ *erzeugten* $\sigma$*-Algebra* $\sigma(\mathcal{R})$ *ist, das* $\nu$ *fortsetzt.*

**Lemma 14.2.11.**

$$\mathcal{F} := \left\{ V = \bigcup_{k=1}^{N} ]a_k, b_k] : N \in \mathbb{N}, a_k, b_k \in \mathbb{R}, a_k \leq b_k \right\} \tag{14.2.15}$$

*ist ein Mengenring auf* $\mathbb{R}$.

**Lemma 14.2.12.** *Es existiert genau ein Inhalt* $\mu$ *auf* $\mathcal{F}$, *so daß*

$$\mu(]a, b]) = b - a \tag{14.2.16}$$

*für alle* $a, b \in \mathbb{R}$ *mit* $a \leq b$ *ist. Für diesen dadurch eindeutig bestimmten Inhalt* $\mu$ *gilt:*

- $\mu(B) \in [0, \infty[$ *für jedes* $B \in \mathcal{F}$.
- $\mu$ *ist ein Prämaß auf* $\mathcal{F}$.

**Satz 14.2.13.** *Es gibt genau ein Borelmaß* $\lambda : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ *mit* $\lambda([a, b]) = b - a$ *für alle* $a, b \in \mathbb{R}$ *mit* $a \leq b$.

*Beweis.* Nach Lemma 14.2.11 ist $\mathcal{F} := \left\{ V = \bigcup_{k=1}^{N} ]a_k, b_k] : N \in \mathbb{N}, a_k, b_k \in \mathbb{R}, a_k \leq b_k \right\}$ ein Mengenring auf $\mathbb{R}$, auf dem es nach Lemma 14.2.12 genau ein Prämaß $\mu$ auf $\mathcal{F}$ mit $\mu(]a, b]) = b - a$ für $a \leq b$ gibt. Nach Satz 14.2.10 gibt es eine Fortsetzung von $\mu$ zu einem Maß $\tilde{\mu}$ auf $\sigma(\mathcal{F})$. Da aber $\mathcal{I} := \{]a, b] : a, b \in \mathbb{R}, a \leq b\} \subseteq \mathcal{F}$, folgt $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{I}) \subseteq \sigma(\mathcal{F})$ nach Lemma 14.1.13, also ist $\lambda := \tilde{\mu}|_{\mathcal{B}(\mathbb{R})}$ ein Maß. Da $\mathcal{I}$ ein durchschnittstabiles Erzeugendensystem von $\mathcal{B}(\mathbb{R})$ ist und zB. $(]-n, n])_{n\in\mathbb{N}}$ eine Folge in $\mathcal{I}$ mit $\lambda(]-n, n]) = 2n < \infty$ und $\mathbb{R} = \bigcup_{n\in\mathbb{N}} ]-n, n]$ ist, folgt die Eindeutigkeit eines Maßes $\lambda : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ mit $\lambda(]a, b]) = b - a$ aus Satz 14.2.8. □

**Bemerkung 14.2.14.** Wegen $\sigma-$Stetigkeit des Maßes $\lambda$ gilt für alle $b \in \mathbb{R}$:

$$\lambda(\{b\}) = \lambda\left(\bigcap_{n\in\mathbb{N}} ]b - \frac{1}{n}, b]\right) = \lim_{n\to\infty} \lambda(]b - \frac{1}{n}, b]) = 0,$$

also folgt auch $\lambda(]a, b]) = \lambda([a, b[) = \lambda([a, b]) = \lambda(]a, b[)$ für alle $a, b \in \mathbb{R}$, $a < b$.

**Bemerkung 14.2.15.** Ist $(X, \mathcal{A}, \mu)$ ein Maßraum, dann definiert

$$\begin{array}{rcl} \mu^* : \mathcal{P}(X) & \to & [0, \infty] \\ Y & \mapsto & \inf\{\mu(A) : A \in \mathcal{A}, A \supseteq Y\} \end{array} \tag{14.2.17}$$

– wie der Beweis von Satz 14.2.10 zeigt – ein äußeres Maß das **äußere Maß zu $\mu$.**

**Definition 14.2.16.** *Ist* $(X, \mathcal{A}, \mu)$ *ein Maßraum und* $\mu^* : \mathcal{P}(X) \to [0, \infty]$ *das äußere Maß zu* $\mu$, *dann heißt jede Menge* $Y \subseteq X$ *mit* $\mu^*(Y) = 0$ *eine* **$\mu$-Nullmenge**. *Eine Familie von Aussagen* $A(x); x \in X$ *gilt* **$\mu$-fast überall**, *wenn* $\{x \in X : A(x)$ *gilt nicht* $\}$ *eine* $\mu$*-Nullmenge ist. Der Maßraum* $(X, \mathcal{A}, \mu)$ *heißt* **vollständig**, *wenn jede* $\mu$*-Nullmenge in* $\mathcal{A}$ *liegt.*

**Lemma 14.2.17.** *Es sei* $(X, \mathcal{A}, \mu)$ *ein Maßraum und* $\mathcal{N} := \{Y \subseteq X : \mu^*(Y) = 0\}$ *dann ist*

$$\tilde{\mathcal{A}} := \{A \cup N : A \in \mathcal{A}, N \in \mathcal{N}\}$$

*eine* $\sigma$*-Algebra auf* $X$ *und durch* $\tilde{\mu}(A \cup N) := \mu(A)$ *wird ein vollständiges Maß* $\tilde{\mu} : \tilde{\mathcal{A}} \to [0, \infty]$ *definiert. Es gilt sogar: Jede Fortsetzung* $\nu : \mathcal{B} \to [0, \infty]$ *von* $\mu$ *zu einem vollständigen Maß* $\nu$ *ist eine Fortsetzung von* $\tilde{\mu}$.

*Beweis.*

- Offenbar ist $X = X \cup \emptyset \in \tilde{\mathcal{A}}$.

- Es sei $M \in \tilde{\mathcal{A}}$, dann gibt es $A \in \mathcal{A}$ und $N \in \mathcal{N}$ mit $M = A \cup N$. Nach Bemerkung 14.2.15 gibt es $C \in \mathcal{A}$ mit $\mu(C) = 0$ und $N \subseteq C$. Dann ist

$$\begin{aligned} X\backslash M &= (X\backslash A) \cap (X\backslash N) = (X\backslash A) \cap ((X\backslash C) \cup (C \cap (X\backslash N))) \\ &= ((X\backslash A) \cap (X\backslash C)) \cup ((X\backslash A) \cap C \cap (X\backslash N) \in \tilde{\mathcal{A}} \end{aligned}$$

denn $(X\backslash A) \cap (X\backslash C) \in \mathcal{A}$ und $(X\backslash A) \cap C \cap (X\backslash N) \subseteq C$ ergibt

$$\mu((X\backslash A) \cap C \cap (X\backslash N)) = 0.$$

- Ist $(M_n)_{n\in\mathbb{N}}$ eine Folge in $\tilde{\mathcal{A}}$, dann gibt es jeweils eine Zerlegung $M_n = A_n \cup N_n$ mit $A_n \in \mathcal{A}$ und $N_n \in \mathcal{N}$, also gibt es auch eine Folge $(C_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ mit $\mu(C_n) = 0$ und $N_n \subseteq C_n$. Dann ist $\bigcup\limits_{n\in\mathbb{N}} M_n = \bigcup\limits_{n\in\mathbb{N}} A_n \cup \bigcup\limits_{n\in\mathbb{N}} N_n \subseteq \bigcup\limits_{n\in\mathbb{N}} A_n \cup \bigcup\limits_{n\in\mathbb{N}} C_n$ und wegen $\sigma$-Subadditivität von $\mu$ ist $\mu\left(\bigcup\limits_{n\in\mathbb{N}} C_n\right) \leq \sum\limits_{n=1}^{\infty} \mu(C_n) = 0$, also $\bigcup\limits_{n\in\mathbb{N}} C_n \supseteq \bigcup\limits_{n\in\mathbb{N}} N_n \in \mathcal{N}$ und damit $\bigcup\limits_{n\in\mathbb{N}} M_n$.

Dies zeigt, daß $\mathcal{A}$ eine $\sigma$-Algebra ist.

- Sind $A, B \in \mathcal{A}$ und $N, Q \in \mathcal{N}$ mit $A \cup N = B \cup Q$, dann gibt es $C \in \mathcal{A}$ mit $N \cup Q \subseteq C$ und $\mu(C) = 0$ und damit ist $A \subseteq B \cup C$, also $\mu(A) \leq \mu(B) + \mu(C) = \mu(B)$. Ebenso ist $B \subseteq A \cup C$ und damit auch $\mu(B) \leq \mu(A)$, dh. $\tilde{\mu}(A \cup N) = \mu(A) = \mu(B) = \tilde{\mu}(B \cup Q)$, also ist $\tilde{\mu}$ wohldefiniert.

- $\tilde{\mu}(\emptyset) = \mu(\emptyset) = 0$

- Ist $(M_n)_{n\in\mathbb{N}}$ eine Folge von paarweise disjunkten Mengen in $\tilde{\mathcal{A}}$, dann gibt es nach Definition von $\tilde{\mathcal{A}}$ Mengen $A_n, C_n \in \mathcal{A}$ und $N_n \in \mathcal{N}$ mit $M_n = A_n \cup N_n$, $\mu(C_n) = 0$ und $N_n \subseteq C_n$. Dann ist auch $(A_n)_{n\in\mathbb{N}}$ eine Folge von paarweise disjunkten Mengen in $\mathcal{A}$ und $\bigcup\limits_{n\in\mathbb{N}} N_n \subseteq \bigcup\limits_{n\in\mathbb{N}} C_n$ eine $\mu$-Nullmenge, also

$$\tilde{\mu}\left(\bigcup_{n\in\mathbb{N}} M_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n) = \sum_{n\in\mathbb{N}} \tilde{\mu}(M_n)$$

$\tilde{\mu}$ ist daher ein Maß auf $\tilde{\mathcal{A}}$. □

**Definition 14.2.18.** *Ist $(X, \mathcal{A}, \mu)$ ein Maßraum so heißt der in Lemma 14.2.17 beschriebene Maßraum $(X, \tilde{\mathcal{A}}, \tilde{\mu})$ die* **Vervollständigung** *von $(X, \mathcal{A}, \mu)$. $\widetilde{\mathcal{B}(\mathbb{R})}$ heißt die* **Lebesguesche $\sigma$-Algebra** *und $\tilde{\lambda} : \widetilde{\mathcal{B}(\mathbb{R})} \to [0, \infty]$ das* **Lebesguemaß** *auf $\mathbb{R}$.*

## 14.3 Integration von nichtnegativen Funktionen

**Lemma 14.3.1.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $u : X \to [0, \infty[$ eine $\mathcal{A}$-Stufenfunktion. Sind $A_1, ..., A_m \in \mathcal{A}$ paarweise disjunkt mit $X = \bigcup_{k=1}^{m} A_k$ und $B_1, ..., B_n \in \mathcal{A}$ paarweise disjunkt mit $X = \bigcup_{l=1}^{n} B_l$ und $\alpha_1, ..., \alpha_m, \beta_1, ..., \beta_n \in [0, \infty[$ mit*

$$u = \sum_{k=1}^{m} \alpha_k \mathbf{1}_{A_k} = \sum_{l=1}^{n} \beta_l \mathbf{1}_{B_l},$$

*dann gilt:*

$$\sum_{k=1}^{m} \alpha_k \mu(A_k) = \sum_{l=1}^{n} \beta_l \mu(B_l).$$

*Beweis.* Wegen $X = A_1 \cup ... \cup A_m = B_1 \cup ... \cup B_n$ ist

$$A_k = \bigcup_{l=1}^{n} (A_k \cap B_l) \quad , \quad B_l = \bigcup_{k=1}^{m} (A_k \cap B_l)$$

mit paarweise disjunkten Mengen $A_k \cap B_l \in \mathcal{A}$, also

$$\mu(A_k) = \sum_{l=1}^{n} \mu(A_k \cap B_l) \quad \text{und} \quad \mu(B_l) = \sum_{k=1}^{m} \mu(A_k \cap B_l).$$

Daher ist

$$\sum_{k=1}^{m} \alpha_k \mu(A_k) = \sum_{k=1}^{m} \sum_{l=1}^{n} \alpha_k \mu(A_k \cap B_l) \quad \text{und} \quad \sum_{l=1}^{n} \beta_l \mu(B_l) = \sum_{l=1}^{n} \sum_{k=1}^{m} \beta_l \mu(A_k \cap B_l)$$

und wegen

$$u = \sum_{k=1}^{m} \alpha_k \mathbf{1}_{A_k} = \sum_{k=1}^{m} \sum_{l=1}^{n} \alpha_k \mathbf{1}_{A_k \cap B_l} = \sum_{l=1}^{n} \beta_l \mathbf{1}_{B_l} = \sum_{l=1}^{n} \sum_{k=1}^{m} \beta_l \mathbf{1}_{A_k \cap B_l}$$

mit paarweise disjunkten Mengen $A_k \cap B_l$ folgt $a_k = \beta_l$ für jedes Tupel $(k, l) \in \{1, ..., m\} \times \{1, ..., n\}$ mit $A_k \cap B_l \neq \emptyset$. $\square$

**Definition 14.3.2.** *Ist $u = \sum_{k=1}^{n} \alpha_k \mathbf{1}_{A_k}$ eine nichtnegative $\mathcal{A}$-Stufenfunktion, so heißt die nach Lemma 14.3.1 eindeutige Zahl*

$$\int_X u d\mu = \int_X u(x) d\mu(x) := \sum_{k=1}^{n} \alpha_k \mu(A_k) \tag{14.3.1}$$

*das* **Integral** *von $u$.*

**Lemma 14.3.3.** *Für das Integral von nichtnegativen $\mathcal{A}$-Stufenfunktionen $u$ und $v$ gilt:*

a) $\int\limits_X \mathbf{1}_A d\mu = \mu(A)$

b) *Für* $\alpha \geq 0$ *ist* $\int\limits_X (\alpha u) d\mu = \alpha \int\limits_X u d\mu$.

c) $\int\limits_X (u+v) d\mu = \int\limits_X u d\mu + \int\limits_X v d\mu$.

d) *Ist* $u(x) \leq v(x)$ *für alle* $x \in X$, *dann gilt* $\int\limits_X u d\mu \leq \int\limits_X v d\mu$

*Beweis.* Tutoriumsblatt □

**Lemma 14.3.4.** *Es sei* $(X, \mathcal{A}, \mu)$ *ein Maßraum und* $(f_n : X \to [0, \infty[)_{n \in \mathbb{N}}$ *eine monoton fallende Folge von nichtnegativen* $\mathcal{A}$*-Stufenfunktionen, die punktweise gegen* $0$ *konvergieren mit* $\int\limits_X f_n d\mu < \infty$ *für jedes* $n \in \mathbb{N}$. *Dann ist* $\lim\limits_{n\to\infty} \int\limits_X f_n d\mu = 0$.

*Beweis.* Wegen der Monotonie des Integral in Lemmma 14.3.3 ist im Fall $\int\limits_X f_1 d\mu = 0$ nichts mehr zu zeigen, daher können wir oE. $\int\limits_X f_1 d\mu \in ]0, \infty[$ voraussetzen.

$$A := \{x \in X : f_1(x) > 0\} = f_1^{-1}(]0, \infty[) \in \mathcal{A}.$$

Da $f_1(X) \subseteq [0, \infty[$ eine endliche Menge ist, existieren $m := \min\{f_1(x) : x \in A\} \in ]0, \infty[$ und $M := \max f_1(X) \in ]0, \infty[$.

$$m\mathbf{1}_A \leq f_1 \tag{14.3.2}$$

also ist $m\mu(A) \leq \int\limits_X f_1 d\mu < \infty$, dh. $\mu(A) < \infty$. Es sei nun $\varepsilon > 0$, dann definiere zu $n \in \mathbb{N}$

$$A_n := \{x \in X : f_n(x) > \varepsilon\} = f_n^{-1}(]\varepsilon, \infty[) \in \mathcal{A}$$

dann ist $A \supseteq A_1 \supseteq A_2 \supseteq ...$, denn die Folge $(f_n)_{n \in \mathbb{N}}$ ist monoton fallend und wegen $\lim\limits_{n\to\infty} f_n(x) = 0$ gilt $\bigcap\limits_{n \in \mathbb{N}} A_n = \{x \in X : \sup\limits_{n \in \mathbb{N}} f_n(x) \geq \varepsilon\} = \emptyset$. Aus der $\emptyset$-Stetigkeit von $\mu$ folgt $\lim\limits_{n\to\infty} \mu(A_n) = 0$ und daher führt die Abschätzung $f_n \leq M\mathbf{1}_{A_n} + \varepsilon \mathbf{1}_{A \setminus A_n}$ dank Monotonie des Integrals zu

$$\int\limits_X f_n d\mu \leq M\mu(A_n) + \varepsilon\mu(A). \tag{14.3.3}$$

Bildet man in (14.3.3) zuerst $\limsup\limits_{n\to\infty}$, dann ist

$$\limsup_{n\to\infty} \int\limits_X f_n d\mu \leq M \limsup_{n\to\infty} \mu(A_n) + \varepsilon\mu(A) = \varepsilon\mu(A),$$

was dann $0 \le \limsup\limits_{n\to\infty} \int\limits_X f_n d\mu \le 0$ im Limes $\varepsilon \searrow 0$ zeigt. $\square$

**Lemma 14.3.5.** *Sind $(f_n)_{n\in\mathbb{N}}$ und $(g_n)_{n\in\mathbb{N}}$ zwei Folgen von nichtnegativen $\mathcal{A}$-Stufenfunktionen mit $f_n(x) \le f_{n+1}(x)$ und $g_n(x) \le g_{n+1}(x)$ für alle $x \in X$ und $n \in \mathbb{N}$ und $f(x) := \lim\limits_{n\to\infty} f_n(x) = \lim\limits_{n\to\infty} g_n(x)$ für jedes $x \in X$ – schreibe dafür kurz $0 \le f_n \nearrow f$ – dann ist*

$$\sup\left\{\int\limits_X f_n d\mu : n \in \mathbb{N}\right\} = \sup\left\{\int\limits_X g_n d\mu : n \in \mathbb{N}\right\} \tag{14.3.4}$$

*Beweis.* Da wir die Rollen von $(f_n)_{n\in\mathbb{N}}$ und $(g_n)_{n\in\mathbb{N}}$ auch vertauschen können, genügt es zu gegebenen $k \in \mathbb{N}$ die Abschätzung

$$\int\limits_X f_k d\mu \le \sup\left\{\int\limits_X g_n d\mu : n \in \mathbb{N}\right\} \tag{14.3.5}$$

zu zeigen. Definiere dazu die $\mathcal{A}$-Stufenfunktionen $\begin{array}{rcl} h_n : X & \to & [0,\infty[ \\ x & \mapsto & \min\{f_k(x), g_n(x)\} \end{array}$. Wegen $f_n(x) \le f(x)$ und $g_n(x) \le g_{n+1}(x) \to f(x)$ für jedes $n \in \mathbb{N}$ und $x \in X$ gilt also $\lim\limits_{n\to\infty} h_n(x) = f_k(x)$ und $f_k(x) \ge f_k(x) - h_n(x) \stackrel{n\to\infty}{\longrightarrow} 0$.

- Im Fall $\int\limits_X f_k d\mu < \infty$ folgt $\int\limits_X (f_k - h_n) d\mu \stackrel{n\to\infty}{\longrightarrow} 0$ nach Lemma 14.3.4, also folgt aus den Rechenregeln in Lemma 14.3.3 für das Integral von $\mathcal{A}$-Stufenfunktionen:

$$\int\limits_X f_k d\mu = \sup\left\{\int\limits_X h_n d\mu : n \in \mathbb{N}\right\} \le \sup\left\{\int\limits_X g_n d\mu : n \in \mathbb{N}\right\}.$$

- Im Fall $\int\limits_X f_k d\mu = \infty$ folgt für $A := \{x \in X : f_k(x) > 0\}$, $\alpha := \min f_k(A) \in ]0,\infty[$ und $\beta := \max f_k(A) \in ]0,\infty[$, daß $\beta\mu(A) \ge \int\limits_X f_k d\mu$ ist, also $\mu(A) = \infty$. Für $B_n := g_n^{-1}(]0,\infty[)$ ist $g_n(B_n)$ eine endliche Menge, daher existiert $\min g_n(B_n) \in ]0,\infty[$ und weil die Folge $(g_n)_{n\in\mathbb{N}}$ monoton steigend ist, gibt es $N \in \mathbb{N}$ mit $\min g_n(B_n) > \frac{\alpha}{2}$ für alle $n \ge N$ – denn sonst gibt es $x \in X$ mit $f(x) = \lim\limits_{n\to\infty} g_n(x) \le \dfrac{\alpha}{2}$ im Widerspruch zu $\alpha \le f_k(x) \le f(x)$. Damit ist $\left(\bigcup\limits_{n\in\mathbb{N}} g_n^{-1}([\frac{\alpha}{2},\infty[)\right) \cap A = A$ und wegen $\sigma$-Stetigkeit von unten ist dann:

$$\int\limits_X g_n d\mu \ge \frac{\alpha}{2}\mu\left(\left\{x \in X : g_n(x) \ge \frac{\alpha}{2}\right\} \cap A\right) \stackrel{n\to\infty}{\longrightarrow} \frac{\alpha}{2}\mu(A)$$

und damit auch $\sup \left\{ \int\limits_X g_n d\mu : n \in \mathbb{N} \right\} = \infty$.

□

**Definition 14.3.6.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $f : X \to [0, \infty]$ sei meßbar. Dann gibt es nach Satz 14.1.22 eine monoton steigende Folge $(f_n)_{n\in\mathbb{N}}$ von nichtnegativen $\mathcal{A}$-Stufenfunktionen mit $0 \leq f_n \nearrow f$. Dann heißt die nach Lemma 14.3.5 eindeutige Zahl*

$$\int\limits_X f d\mu := \sup \left\{ \int\limits_X f_n d\mu : n \in \mathbb{N} \right\} = \lim_{n\to\infty} \left\{ \int\limits_X f_n d\mu \right\} \in [0, \infty] \tag{14.3.6}$$

*das* **Integral der nichtnegativen meßbaren Funktion** *$f$.*

**Lemma 14.3.7.** *Für das Integral von nichtnegativen $\mathcal{A}$-meßbaren Funktionen $f$ und $g$ gilt:*

*a) Für $\alpha \geq 0$ ist $\int\limits_X (\alpha f) d\mu = \alpha \int\limits_X f d\mu$.*

*b) $\int\limits_X (f + g) d\mu = \int\limits_X f d\mu + \int\limits_X g d\mu$.*

*c) Ist $f(x) \leq g(x)$ für $\mu$-fast alle $x \in X$, dann gilt $\int\limits_X f d\mu \leq \int\limits_X g d\mu$*

*Beweis.* Wähle Folgen $(f_n)_{n\in\mathbb{N}}$ und $(g_n)_{n\in\mathbb{N}}$ von nichtnegativen $\mathcal{A}$-Stufenfunktionen mit $f_n \nearrow f$ und $g_n \nearrow g$. Dann gilt nach Definition des Integrals bzw. nach Lemma 14.3.3

$$\int\limits_X \alpha f d\mu = \lim_{n\to\infty} \int\limits_X \alpha f_n d\mu = \lim_{n\to\infty} \alpha \int\limits_X f_n d\mu = \alpha \int\limits_X f d\mu$$

$$\begin{aligned} \int\limits_X (f + g) d\mu &= \lim_{n\to\infty} \int\limits_X (f_n + g_n) d\mu = \lim_{n\to\infty} \left( \int\limits_X f_n d\mu + \int\limits_X g_n d\mu \right) \\ &= \lim_{n\to\infty} \int\limits_X f_n d\mu + \lim_{n\to\infty} \int\limits_X g_n d\mu = \int\limits_X f d\mu + \int\limits_X g d\mu. \end{aligned}$$

Ist $f \leq g$, dann definiert $h_n := \min\{f_n, g_n\}$ wieder eine $\mathcal{A}$-Stufenfunktion mit $h_n \nearrow f$ und $h_n \leq g_n$. Aus der Monotonie des $\mathcal{A}$-Stufenfunktionen Integrals und der Definition des Integrals nichtnegativer meßbarer Funktionen folgt:

$$\int\limits_X f d\mu \xleftarrow{n\to\infty} \int\limits_X h_n d\mu \leq \int\limits_X g_n d\mu \xrightarrow{n\to\infty} \int\limits_X g d\mu$$

□

**Satz 14.3.8** (monotone Konvergenz)**.** *Es sei* $(X, \mathcal{A}, \mu)$ *ein Maßraum und* $(f_n)_{n\in\mathbb{N}}$ *eine monoton steigende Folge von nichtnegativen meßbaren Funktionen und*

$$\begin{array}{rcl} f : X & \to & [0,\infty] \\ x & \mapsto & \lim\limits_{n\to\infty} f_n(x) = \sup\limits_{n\in\mathbb{N}} f_n(x) \end{array} .$$

*Dann ist* $f$ *eine nichtnegative* $\mathcal{A}$*-meßbare Funktion und*

$$\int\limits_X f d\mu = \lim_{n\to\infty} \left\{ \int\limits_X f_n d\mu \right\} \tag{14.3.7}$$

*Beweis.* Wie bei der Definition des Integrals können wir für jede nichtnegative $\mathcal{A}$-meßbare Funktion $f_n : X \to [0,\infty]$ eine Folge $(f_{n,k} : X \to [0,\infty[)_{k\in\mathbb{N}}$ von nichtnegativen $\mathcal{A}$-Stufenfunktionen mit $0 \leq f_{n,k} \leq f_{n,k+1} \nearrow f_n$ wählen. Dann ist durch

$$g_n(x) := \sup\{f_{l,k}(x) : 1 \leq k, l \leq n\}$$

eine nichtnegative $\mathcal{A}$-Stufenfunktion $g_n : X \to [0,\infty[$ definiert. Es gilt:

- Für alle $x \in X$ und $n \in \mathbb{N}$ ist
$$0 \leq g_n(x) \leq f_n(x) \leq f(x). \tag{14.3.8}$$
- $(g_n)_{n\in\mathbb{N}}$ ist monoton steigend mit
$$\begin{aligned} \lim_{n\to\infty} g_n(x) &= \sup\{g_n(x) : n \in \mathbb{N}\} = \sup\{f_{l,k}(x) : k, l \in \mathbb{N}\} \\ &= \sup\{f_n(x) : n \in \mathbb{N}\} = f(x) \end{aligned} \tag{14.3.9}$$

Nach Definition des Integrals gilt also

$$\int\limits_X f d\mu = \sup \left\{ \int\limits_X g_n d\mu : n \in \mathbb{N} \right\} = \lim_{n\to\infty} \int\limits_X g_n d\mu \tag{14.3.10}$$

und aus (14.3.8) folgt wegen Monotonie des Integrals

$$0 \leq \int\limits_X g_n d\mu \leq \int\limits_X f_n d\mu \leq \int\limits_X f d\mu, \tag{14.3.11}$$

dh. aus (14.3.9) bis (14.3.11) folgt:

$$\lim_{n\to\infty} \int\limits_X g_n d\mu = \lim_{n\to\infty} \int\limits_X f_n d\mu = \int\limits_X f d\mu.$$

□

**Korollar 14.3.9** (Fatou)**.** *Ist* $(X, \mathcal{A}, \mu)$ *ein Maßraum, so gilt für jede Folge* $(f_n)_{n\in\mathbb{N}}$ *von nichtnegativen* $\mathcal{A}$*-meßbaren Funktionen*

$$\int\limits_X \left( \liminf_{n\to\infty} f_n \right) d\mu \leq \liminf_{n\to\infty} \left\{ \int\limits_X f_n d\mu \right\}. \tag{14.3.12}$$

*Beweis.* Nach Korollar 14.1.17 sind $f := \liminf_{n\to\infty} f_n$ und $g_n := \inf\{f_k : k \geq n\}$ nichtnegative $\mathcal{A}$-meßbare Funktionen und laut Definition des Limesinferiors ist $g_n \nearrow f$. Aus dem Satz über monotone Konvergenz folgt also

$$\int_X (\liminf_{n\to\infty} f_n) d\mu = \int_X f d\mu = \lim_{n\to\infty} \int_X g_n d\mu.$$

Andererseits ist $g_n(x) \leq f_k(x)$ für alle $k \geq n$ und daher wegen Monotonie des Integrals

$$\int_X g_n d\mu \leq \int_X f_k d\mu \qquad \text{für alle } k \geq n$$

also nach Infimumsbildung

$$\int_X g_n d\mu \leq \inf \left\{ \int_X f_k d\mu : k \geq n \right\} \overset{n\to\infty}{\longrightarrow} \liminf_{n\to\infty} \int_X f_n d\mu.$$ □

**Bemerkung 14.3.10.** Wie etwa das Beispiel $(f_n)_{n\in\mathbb{N}} = (n\mathbf{1}_{]0,\frac{1}{n}]})_{n\in\mathbb{N}}$ zeigt, was punktweise gegen 0 konvergiert, also $f_n(x) \overset{n\to\infty}{\longrightarrow} 0$ für jedes $x \in \mathbb{R}$ – aber nicht monoton – kann in Satz 14.3.8 auf die Monotonie der Funktionenfolge nicht verzichtet werden.

**Lemma 14.3.11.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $f : X \to [0, \infty]$ $\mathcal{A}$-meßbar. Dann gilt:*

*a) Ist $\int_X f d\mu < \infty$, dann ist $f(x) < \infty$ $\mu$-fast überall.*

*b) Es ist $\int_X f d\mu = 0$ genau dann wenn $f(x) = 0$ $\mu$-fast überall.*

*Beweis.*

a) Ist $\int_X f d\mu < \infty$ und $A := \{x \in X : f(x) = \infty\} = f^{-1}(\{\infty\}) \in \mathcal{A}$, dann gilt $\mathbf{1}_A(x) \leq \frac{1}{n} f(x)$ für jedes $n \in \mathbb{N}$. Damit ist

$$0 \leq \mu^*(A) = \mu(A) \leq \frac{1}{n} \int_X f d\mu \overset{n\to\infty}{\longrightarrow} 0.$$

b) Betrachte

$$B := \{x \in X : f(x) > 0\} = f^{-1}(]0, \infty]) \in \mathcal{A}$$

und

$$B_n := \{x \in X : f(x) \geq \frac{1}{n}\} = f^{-1}([\frac{1}{n}, \infty]) \in \mathcal{A},$$

so ist $B_1 \subseteq B_2 \subseteq ...$ und $B = \bigcup_{n\in\mathbb{N}} B_n$.

„⇒“: Es sei $\int\limits_X f d\mu = 0$, so folgt aus $\frac{1}{n}\mathbf{1}_{B_n} \leq f$ die Abschätzung

$$0 \leq \frac{1}{n}\mu(B_n) \leq \int\limits_X f d\mu = 0.$$

Daraus folgt $\mu(B_n) = 0$ für jedes $n \in \mathbb{N}$ und damit auch $\mu(B) = 0$ aufgrund der $\sigma$-Stetigkeit von unten von $\mu$.

„⇐“: Es sei $\mu(B) = 0$, dann ist $f_n := n\mathbf{1}_B$ für jedes $n \in \mathbb{N}$ eine nichtnegative $\mathcal{A}$-Stufenfunktion mit $\int\limits_X f_n d\mu = n\mu(B) = 0$. Nach Korollar 14.1.17 ist

$$\begin{aligned} g := \sup\{f_n : n \in \mathbb{N}\} : X &\to [0, \infty] \\ x &\mapsto \sup\{f_n(x) : n \in \mathbb{N}\} \end{aligned}$$

$\mathcal{A}$-meßbar, $f \leq g$ und $0 \leq f_n \leq f_{n+1} \nearrow g$, daher ist

$$\int\limits_X g d\mu = \sup\left\{\int\limits_X f_n d\mu\right\} = 0$$

und aus Monotonie des Integrals folgt $0 \leq \int\limits_X f d\mu \leq \int\limits_X g d\mu = 0$.

□

## 14.4 Integrierbare komplexwertige Funktionen

**Bemerkung 14.4.1.** Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $f : X \to \widehat{\mathbb{R}}$, $g : X \to \mathbb{C}$ seien $\mathcal{A}$-meßbar. Dann sind $|f|$, $f_+$ und $f_-$ nach Korollar 14.1.19 meßbar, daher sind

$$\int\limits_X |f| d\mu, \int\limits_X f_+ d\mu, \int\limits_X f_- d\mu \in [0, \infty]$$

in Kapitel 14.3 definiert worden. Ist nun

$$\int\limits_X |f| d\mu < \infty,$$

dann folgt wegen $f_+, f_- \leq f_+ + f_- = |f|$ aus der Monotonie des Integrals nichtnegativer meßbarer Funktionen auch

$$\int\limits_X f_+ d\mu, \int\limits_X f_- d\mu < \infty$$

und damit ist in diesem Fall

$$\int\limits_X f d\mu := \int\limits_X f_+ d\mu - \int\limits_X f_- d\mu \in \mathbb{R}$$

wohldefiniert. Die Abbildungen $\operatorname{Re} g : X \to \mathbb{R}$ und $\operatorname{Im} g : X \to \mathbb{R}$ und damit auch $(\operatorname{Re} g)_+, (\operatorname{Re} g)_-, (\operatorname{Im} g)_+, (\operatorname{Im} g)_-$ sind $\mathcal{A}$-meßbar und $|g(x)| \geq |\operatorname{Re} g(x)|, |\operatorname{Im} g(x)|$. Ist also $\int\limits_X |g| d\mu < \infty$, dann ist auch $\int\limits_X (\operatorname{Re} g)_+ d\mu, \int\limits_X (\operatorname{Re} g)_- d\mu, \int\limits_X (\operatorname{Im} g)_+ d\mu, \int\limits_X (\operatorname{Im} g)_- d\mu \in [0, \infty[$, also

$$\begin{aligned}\int\limits_X g d\mu &:= \int\limits_X (\operatorname{Re} g) d\mu + i \int\limits_X (\operatorname{Im} g) d\mu \\ &= \int\limits_X (\operatorname{Re} g)_+ d\mu - \int\limits_X (\operatorname{Re} g)_- d\mu + i \int\limits_X (\operatorname{Im} g)_+ d\mu - i \int\limits_X (\operatorname{Im} g)_- d\mu \in \mathbb{C}\end{aligned}$$

wohldefiniert.

**Definition 14.4.2.** *Es sei* $(X, \mathcal{A}, \mu)$ *ein Maßraum.* $f : X \to \mathbb{C}$ *heißt* $\mu$-**integrierbar**, *wenn* $f$ $\mathcal{A}$-*meßbar ist und*

$$\int\limits_X |f(x)| d\mu(x) < \infty \tag{14.4.1}$$

*gilt. In diesem Fall heißt*

$$\int\limits_X f d\mu := \int\limits_X (\operatorname{Re} f)_+ d\mu - \int\limits_X (\operatorname{Re} f)_- d\mu + i \int\limits_X (\operatorname{Im} f)_+ d\mu - i \int\limits_X (\operatorname{Im} f)_- d\mu \tag{14.4.2}$$

*das* **Integral von** $f$ *und für* $A \in \mathcal{A}$ *heißt*

$$\int\limits_A f d\mu := \int\limits_X (\mathbf{1}_A \cdot f) d\mu \tag{14.4.3}$$

*das* **Integral von** $f$ **über** $A$.

**Bemerkung 14.4.3.** Bei einer $\lambda-$integrierbaren Funktion $f : \mathbb{R} \to \mathbb{R}$, wird beim in Definition 14.4.2 definiertem Integral $\int\limits_{\mathbb{R}} f d\lambda$ die Fläche in der oberen Halbebene $\mathbb{R} \times [0, \infty[$ zwischen reeller Achse und Graph – also das Integral $\int\limits_{\mathbb{R}} f_+ d\lambda$ über den Positivteil von $f$ – auch positiv gezählt. Die Fläche in der unteren Halbebene $\mathbb{R} \times ]-\infty, 0]$ zwischen Graph und reeller Achse – also das Integral $\int\limits_{\mathbb{R}} f_- d\lambda$ über den Negativteil von $f$ – wird negativ gezählt.

**Lemma 14.4.4.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum.*

*a) Sind $f : X \to \mathbb{R}$ und $g : X \to \mathbb{R}$ $\mu$-intergrierbar mit $f(x) \leq g(x)$ für alle $x \in X$, dann gilt auch $\int\limits_X f d\mu \leq \int\limits_X g d\mu$.*

*b) Sind $f : X \to \mathbb{C}$ und $g : X \to \mathbb{C}$ $\mu$-intergrierbar und $a \in \mathbb{C}$, dann ist auch $af + g$ $\mu$-integrierbar mit*

$$\int\limits_X (af + g) d\mu = a \int\limits_X f d\mu + \int\limits_X g d\mu \tag{14.4.4}$$

*c) Für jede meßbare Funktion $f : X \to \mathbb{C}$ und jedes $\alpha \in \mathbb{R}$ gilt die* **Tschebyscheff-Ungleichung**

$$\alpha\mu\left(\{x \in X : |f(x)| \geq \alpha\}\right) \leq \int\limits_X |f| d\mu. \tag{14.4.5}$$

*d) Ist $f : X \to \mathbb{C}$ $\mu$-integrierbar, so gilt:*

$$\left| \int\limits_X f d\mu \right| \leq \int\limits_X |f| d\mu. \tag{14.4.6}$$

*Beweis.*

a) Sind $f : X \to \mathbb{R}$ und $g : X \to \mathbb{R}$ mit $f \leq g$, dann ist $f_+ \leq g_+$ und $f_- \geq g_-$, also

$$\int\limits_X f d\mu = \int\limits_X f_+ d\mu - \int\limits_X f_- d\mu \leq \int\limits_X g_+ d\mu - \int\limits_X g_- d\mu = \int\limits_X g d\mu$$

wegen Monotonie des Integrals nichtnegativer $\mathcal{A}$-meßbarer Funktionen.

b) Offenbar gilt $\mathrm{Re}\,(f + g) = \mathrm{Re}\, f + \mathrm{Re}\, g$, $\mathrm{Im}\,(f + g) = \mathrm{Im}\, f + \mathrm{Im}\, g$ und $\mathrm{Re}\,(af) = \alpha \mathrm{Re}\, f - \beta \mathrm{Im}\, f$, $\mathrm{Im}\,(af) = \alpha \mathrm{Im}\, f + \beta \mathrm{Re}\, f$ für meßbare Funktionen $f : X \to \mathbb{C}$ und $g : X \to \mathbb{C}$ und $a = \alpha + i\beta \in \mathbb{C}$. Wegen $|af(x) + g(x)| \leq |a||f(x)| + |g(x)|$ ist

$$\int\limits_X |af + g| d\mu \leq |a| \int\limits_X |f| d\mu + \int\limits_X |g| d\mu < \infty$$

also $af + g$ $\mu$-integrierbar und

$$\begin{aligned}
\int\limits_X (f + g) d\mu &= \int\limits_X (\mathrm{Re}\, f + \mathrm{Re}\, g) d\mu + i \int\limits_X (\mathrm{Im}\, f + \mathrm{Im}\, g) d\mu \\
\int\limits_X (\alpha + i\beta) f d\mu &= \int\limits_X (\alpha \mathrm{Re}\, f - \beta \mathrm{Im}\, f) d\mu + i \int\limits_X (\alpha \mathrm{Im}\, f + \beta \mathrm{Re}\, f) d\mu.
\end{aligned}$$

Daher genügt es meßbare Funktionen $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}$ und $a = \alpha \in \mathbb{R}$ zu betrachten und wir zerlegen $\alpha = \alpha_+ - \alpha_- \in \mathbb{R}$. Ferner ist $(\alpha f)_+ = \alpha_+ f_+ + \alpha_- f_-$ und $(\alpha f)_- = \alpha_- f_+ + \alpha_+ f_-$, also

$$\begin{aligned}
\int_X \alpha f d\mu &:= \int_X (\alpha_+ f_+ + \alpha_- f_-) d\mu - \int_X (\alpha_- f_+ + \alpha_+ f_-) d\mu \\
&= \alpha_+ \int_X f_+ d\mu + \alpha_- \int_X f_- d\mu - \alpha_- \int_X f_+ d\mu - \alpha_+ \int_X f_- d\mu \\
&= (\alpha_+ - \alpha_-) \int_X f_+ d\mu - (\alpha_+ - \alpha_-) \int_X f_- d\mu =: \alpha \int_X f d\mu
\end{aligned}$$

nach Definition des Integrals und der Rechenregeln für das Integral von nichtnegativen meßbaren Funktionen in Lemma 14.3.7. Wegen

$$\begin{aligned}
(f+g)_+ \mathbf{1}_{\{f_+>0, g_+>0\}} &= (f_+ + g_+) \mathbf{1}_{\{f_+>0, g_+>0\}} \\
(f+g)_+ \mathbf{1}_{\{f_+=0, g_+=0\}} &= 0 \\
(f+g)_+ \mathbf{1}_{\{f_+>0, g_+=0\}} &= (f_+ - g_-) \mathbf{1}_{\{f_+>0, g_+=0\}} \mathbf{1}_{\{f_+ \geq g_-\}} \\
(f+g)_+ \mathbf{1}_{\{f_+=0, g_+>0\}} &= (g_+ - f_-) \mathbf{1}_{\{f_+=0, g_+>0\}} \mathbf{1}_{\{g_+ \geq f_-\}} \\
(f+g)_- \mathbf{1}_{\{f_->0, g_->0\}} &= (f_- + g_-) \mathbf{1}_{\{f_->0, g_->0\}} \\
(f+g)_- \mathbf{1}_{\{f_-=0, g_-=0\}} &= 0 \\
(f+g)_- \mathbf{1}_{\{f_->0, g_-=0\}} &= (f_- - g_+) \mathbf{1}_{\{f_->0, g_-=0\}} \mathbf{1}_{\{f_- > g_+\}} \\
(f+g)_- \mathbf{1}_{\{f_-=0, g_->0\}} &= (g_- - f_+) \mathbf{1}_{\{f_-=0, g_->0\}} \mathbf{1}_{\{g_- > f_+\}}
\end{aligned}$$

ist

$$\begin{aligned}
&\int_X (f+g) d\mu = \\
&= \int_X f_+ (\mathbf{1}_{\{f_+>0, g_+>0\}} + \mathbf{1}_{\{f_+>0, g_+=0\}} \mathbf{1}_{\{f_+ \geq g_-\}} + \mathbf{1}_{\{f_+>0, g_+=0\}} \mathbf{1}_{\{f_+ < g_-\}}) d\mu \\
&\quad + \int_X g_+ (\mathbf{1}_{\{f_+>0, g_+>0\}} + \mathbf{1}_{\{f_+=0, g_+>0\}} \mathbf{1}_{\{g_+ \geq f_-\}} + \mathbf{1}_{\{f_->0, g_-=0\}} \mathbf{1}_{\{f_- > g_+\}}) d\mu \\
&\quad - \int_X f_- (\mathbf{1}_{\{f_->0, g_->0\}} + \mathbf{1}_{\{f_+=0, g_+>0\}} \mathbf{1}_{\{g_+ \geq f_-\}} + \mathbf{1}_{\{f_->0, g_-=0\}} \mathbf{1}_{\{f_- > g_+\}}) d\mu \\
&\quad - \int_X g_- (\mathbf{1}_{\{f_->0, g_->0\}} + \mathbf{1}_{\{f_+>0, g_+>0\}} \mathbf{1}_{\{f_+ \geq g_-\}} + \mathbf{1}_{\{f_+>0, g_+=0\}} \mathbf{1}_{\{f_+ < g_-\}}) d\mu \\
&= \int_X f_+ \mathbf{1}_{\{f_+>0\}} d\mu + \int_X g_+ \mathbf{1}_{\{g_+>0\}} d\mu - \int_X f_- \mathbf{1}_{\{f_->0\}} d\mu - \int_X g_- \mathbf{1}_{\{g_->0\}} d\mu \\
&= \int_X f_+ d\mu + \int_X g_+ d\mu - \int_X f_- d\mu - \int_X g_- d\mu = \int_X f d\mu + \int_X g d\mu
\end{aligned}$$

c) Für jedes $\alpha \in \mathbb{R}$ gilt (ganz unabhängig vom Vorzeichen von $\alpha$):

$$\alpha \mathbf{1}_{\{x \in X : |f(x)| \geq \alpha\}}(x) \leq |f(x)|. \tag{14.4.7}$$

Integriert man (14.4.7), so folgt (14.4.5).

d) Wir dürfen zum Beweis von (14.4.6) oE. $\int\limits_X f d\mu \neq 0$ voraussetzen, dann ist

$$\alpha := \frac{\left|\int\limits_X f d\mu\right|}{\int\limits_X f d\mu} \in \mathbb{C} \text{ mit } |\alpha| = 1.$$

Nach Wahl von $\alpha$ gilt

$$\left|\int\limits_X f d\mu\right| = \alpha \int\limits_X f d\mu = \int\limits_X (\alpha f) d\mu.$$

Da aber $\left|\int\limits_X f d\mu\right| \in \mathbb{R}$ ist, folgt

$$\left|\int\limits_X f d\mu\right| = \operatorname{Re} \int\limits_X (\alpha f) d\mu = \int\limits_X \operatorname{Re}(\alpha f) d\mu \leq \int\limits_X |\alpha f| d\mu = \int\limits_X |f| d\mu.$$

□

**Satz 14.4.5.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $f : X \to \mathbb{C}$ sei $\mu$-integrierbar. Dann sind äquivalent:*

*a) $f(x) = 0$ $\mu$-fast überall*

*b) Für jedes $A \in \mathcal{A}$ gilt $\int\limits_A f d\mu = 0$.*

*Beweis.* Wegen $f(x) = (\operatorname{Re} f)_+(x) - (\operatorname{Re} f)_-(x) + i(\operatorname{Im} f)_+(x) - i(\operatorname{Im} f)_-(x)$ ist $f = 0$ $\mu$-fast überall genau dann wenn $(\operatorname{Re} f)_+ = 0$, $(\operatorname{Re} f)_-, = 0$, $(\operatorname{Im} f)_+ = 0$ und $(\operatorname{Im} f)_- = 0$ $\mu$-fast überall

a)⇒b) Ist $f = 0$ $\mu$-fast überall, dann ist auch $(\operatorname{Re} f)_+ = 0$ $\mu$-fast überall, also ist nach Lemma 14.3.11

$$0 \leq \int\limits_A (\operatorname{Re} f)_+ d\mu \leq \int\limits_X (\operatorname{Re} f)_+ d\mu = 0$$

und analog

$$\int_A (\operatorname{Re} f)_- d\mu = \int_A (\operatorname{Im} f)_+ d\mu = \int_A (\operatorname{Im} f)_- d\mu = 0,$$

also $\int_A f d\mu = 0$.

b)⇒a) Ist $\int_A f d\mu = 0$ für jedes $A \in \mathcal{A}$, so folgt $\mathbf{1}_A \operatorname{Re} f = (\operatorname{Re} f)_+$ für

$$A := \{x \in X : (\operatorname{Re} f)(x) \geq 0\} \in \mathcal{A}$$

und daher nach Voraussetzung

$$0 = \int_A f d\mu = \int_X (\operatorname{Re} f)_+ d\mu + i \int_A (\operatorname{Im} f) d\mu,$$

also $\int_X (\operatorname{Re} f)_+ d\mu = 0$ und somit $(\operatorname{Re} f)_+ = 0$ $\mu$-fast überall nach Lemma 14.3.11. Analog zeigt man mit $\{x \in X : \operatorname{Re} f(x) < 0\}$ die Gleichheit $(\operatorname{Re} f)_- = 0$ $\mu$-fast überall und ebenso geht es mit dem Imaginärteil. □

**Satz 14.4.6** (majorisierte Konvergenz)**.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum, $(f_n : X \to \mathbb{C})_{n\in\mathbb{N}}$ eine Folge von $\mathcal{A}$ meßbaren Funktionen und $h : X \to [0, \infty]$ sei $\mu$-integrierbar mit $|f_n(x)| \leq h(x)$ für alle $x \in X$. ( Sprechweise: $h$* **integrierbare Majorante***.) Gilt $f_n \overset{n\to\infty}{\longrightarrow} f$ punktweise, so sind $f$ und alle $f_n$ $\mu$-integrierbar und es gilt:*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu. \tag{14.4.8}$$

*Beweis.* Nach Korollar 14.1.17 ist $f$ als punktweiser Grenzwert von meßbaren Funktionen wieder meßbar. Aus $|f_n(x)| \leq h(x)$ erhält man für $n \to \infty$ die Abschätzung $|f(x)| \leq h(x)$ und damit ist

$$\int_X |f| d\mu \leq \int_X |h| d\mu < \infty,$$

also ist $f$ $\mu$-integrierbar. Wegen $f(x) = \lim_{n\to\infty} f_n(x)$ ist auch

$$(\operatorname{Re} f)_+(x) = \lim_{n\to\infty} (\operatorname{Re} f_n)_+(x) = \liminf_{n\to\infty} (\operatorname{Re} f)_+(x) = \limsup_{n\to\infty} (\operatorname{Re} f)_+(x)$$

und nach dem Lemma von Fatou folgt:

$$\liminf_{n\to\infty} \int_X (\operatorname{Re} f_n)_+ d\mu \geq \int_X \liminf_{n\to\infty} (\operatorname{Re} f_n)_+ d\mu = \int_X (\operatorname{Re} f)_+ d\mu. \tag{14.4.9}$$

Wegen $(\operatorname{Re} f_n)_+ \leq |\operatorname{Re} f_n| \leq |f_n| \leq h$ ist auch $h - (\operatorname{Re} f_n)_+$ eine nichtnegative $\mathcal{A}$-meßbare Funktion, daher ergibt das Lemma von Fatou

$$\begin{aligned}\liminf_{n\to\infty} \int\limits_X (h - (\operatorname{Re} f_n)_+) d\mu &\geq \int\limits_X \liminf_{n\to\infty} (h - (\operatorname{Re} f_n)_+) d\mu = \int\limits_X (h - (\operatorname{Re} f)_+) d\mu \\ &= \int\limits_X h d\mu - \int\limits_X (\operatorname{Re} f)_+ d\mu \end{aligned} \tag{14.4.10}$$

Nach den Rechenregeln für lim inf und lim sup ist

$$\liminf_{n\to\infty} \int\limits_X (h - (\operatorname{Re} f_n)_+) d\mu = \int\limits_X h d\mu - \limsup_{n\to\infty} \int\limits_X (\operatorname{Re} f_n)_+ d\mu \geq \int\limits_X h d\mu - \int\limits_X (\operatorname{Re} f)_+ d\mu$$

und daher folgt aus (14.4.9) und (14.4.10):

$$\liminf_{n\to\infty} \int\limits_X (\operatorname{Re} f_n)_+ d\mu \leq \limsup_{n\to\infty} \int\limits_X (\operatorname{Re} f_n)_+ d\mu \leq \int\limits_X (\operatorname{Re} f)_+ d\mu \leq \liminf_{n\to\infty} \int\limits_X (\operatorname{Re} f_n)_+ d\mu$$

also die Gleichheit

$$\int\limits_X (\operatorname{Re} f)_+ d\mu = \lim_{n\to\infty} \int\limits_X (\operatorname{Re} f_n)_+ d\mu. \tag{14.4.11}$$

Der Nachweis für $(\operatorname{Re} f)_-$ bzw. für den Imaginärteil geht analog. □

**Korollar 14.4.7.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum, $Y$ ein metrischer Raum. Für $f : X \times Y \to \mathbb{C}$ sei zu jedem $y \in Y$ die Funktion $\begin{array}{rcl} f(\cdot, y) : X & \to & \mathbb{C} \\ x & \mapsto & f(x, y) \end{array}$ $\mu$-integrierbar, also existiert*

$$\begin{array}{rcl} g : Y & \to & \mathbb{C} \\ y & \mapsto & \int\limits_X f(\cdot, y) d\mu = \int\limits_X f(x, y) d\mu(x) \end{array}.$$

*Für alle $x \in X$ sei $\begin{array}{rcl} f(x, \cdot) : Y & \to & \mathbb{C} \\ y & \mapsto & f(x, y) \end{array}$ stetig und gebe es eine $\mu$-integrierbare Funktion $h : X \to [0, \infty]$ mit $|f(x, y)| \leq h(x)$ für alle $(x, y) \in X \times Y$, dann ist $g$ stetig.*

*Beweis.* Unter diesen Voraussetzungen existiert $g : Y \to \mathbb{C}$. Wir wählen eine Folge $(a_n)_{n\in\mathbb{N}}$ in $Y$ mit $\lim\limits_{n\to\infty} a_n = a$ und setzen $f_n := f(\cdot, a_n)$, so folgt aus der Stetigkeit von $f(x, \cdot)$ die punktweise Konvergenz $f_n(x) \to f(x, a)$ für jedes $x \in X$. Wir können jetzt den Satz von der majorisierten Konvergenz auf die Folge $(f_n)_{n\in\mathbb{N}}$ anwenden und bekommen

$$\lim_{n\to\infty} g(a_n) = \lim_{n\to\infty} \int\limits_X f_n d\mu = \int\limits_X \left(\lim_{n\to\infty} f_n\right) d\mu = \int\limits_X f(\cdot, a) d\mu = g(a)$$

□

**Lemma 14.4.8.** *Ist* $(X, \mathcal{A}, \mu)$ *ein Maßraum,* $(Y, \mathcal{B})$ *ein Meßraum,* $f : X \to Y$ *und* $g : Y \to \mathbb{C}$ *meßbar, dann sind äquivalent:*

*a)* $g$ *ist* $\mu \circ f^{-1}$*-integrierbar*

*b)* $g \circ f$ *ist* $\mu$*-integrierbar*

*und in diesem Fall gilt:*

$$\int\limits_X (g \circ f) d\mu = \int\limits_Y g d(\mu \circ f^{-1}) \tag{14.4.12}$$

*Beweis.* Nach Voraussetzung sind $f$ und $g$ meßbar und damit ist auch $g \circ f$ meßbar. Ist $A \in \mathcal{B}$ und $g = \mathbf{1}_A$, dann ist

$$\int\limits_X |g \circ f| d\mu = \int\limits_X \mathbf{1}_A(f(x)) d\mu(x) = (\mu \circ f^{-1})(A) = \int\limits_X |g| d(\mu \circ f^{-1}). \tag{14.4.13}$$

Sind nun $\alpha_1, ..., \alpha_n \in [0, \infty[$ und $B_1, ..., B_n \in \mathcal{B}$ und ist $g = \sum\limits_{k=1}^{n} \alpha_k \mathbf{1}_{B_k}$ eine nichtnegative $\mathcal{B}-$Stufenfunktion, dann ist

$$\begin{aligned}\int\limits_X |g \circ f| d\mu &= \int\limits_X \left| \left( \sum_{k=1}^{n} \alpha_k \mathbf{1}_{B_k} \right) \circ f \right| d\mu = \sum_{k=1}^{n} \alpha_k \int\limits_X \mathbf{1}_{B_k} \circ f d\mu = \sum_{k=1}^{n} \alpha_k (\mu \circ f^{-1}(A_k) \\ &= \int\limits_Y \sum_{k=1}^{n} \alpha_k \mathbf{1}_{B_k} d(\mu \circ f^{-1}) = \int\limits_Y |g| d(\mu \circ f^{-1}),\end{aligned} \tag{14.4.14}$$

was die Gültigkeit von (14.4.13) von charakteristischen Funktionen $g = \mathbf{1}_A$ auf nichtnegative $\mathcal{B}-$Stufenfunktionen erweitert. Wähle nun eine monoton steigende Folge $(g_n : Y \to [0, \infty[)_{n \in \mathbb{N}}$ von $\mathcal{B}-$Stufenfunktionen mit $0 \leq g_n \nearrow |g|$, dann folgt

$$\int\limits_X |g \circ f| d\mu = \lim_{n \to \infty} \int\limits_X g_n \circ f d\mu = \lim_{n \to \infty} \int\limits_X g_n d(\mu \circ f^{-1}) = \int\limits_Y |g| d(\mu \circ f^{-1}). \tag{14.4.15}$$

Wegen (14.4.15) ist eine meßbare Funktion $g : Y \to \mathbb{C}$ genau dann $\mu \circ f^{-1}-$integrierbar, wenn $g \circ f$ $\mu-$integrierbar ist. Durch Approximation von $(\operatorname{Re} g)_+$, $(\operatorname{Re} g)_-$, $(\operatorname{Im} g)_+$ bzw. $(\operatorname{Im} g)_-$ durch jeweils eine monoton steigende Folge von nichtnegativen Stufenfunktionen erhält man analog zu (14.4.15) auch

$$\int\limits_X (\operatorname{Re} g)_\pm \circ f d\mu = \int\limits_Y (\operatorname{Re} g)_\pm d(\mu \circ f^{-1}) \text{ und } \int\limits_X (\operatorname{Im} g)_\pm \circ f d\mu = \int\limits_Y (\operatorname{Im} g)_\pm d(\mu \circ f^{-1})$$

also im Fall $\int\limits_X |g \circ f| d\mu = \int\limits_Y |g| d(\mu \circ f^{-1} < \infty$ auch

$$\int\limits_X g \circ f d\mu = \int\limits_Y g d(\mu \circ f^{-1})$$

□

## 14.5 Ein Crashkurs im Integrieren von Banachraum-wertigen Funktionen

**Definition 14.5.1.** *Es sei $(X, \mathcal{A}, \mu)$ ein vollständiger, $\sigma$-endlicher Maßraum und $Y$ ein Banachraum. Eine Funktion $f : X \to Y$ heißt* $\mu$**-einfach**, *wenn*

- *$f(X)$ eine endliche Menge ist.*
- *Für jedes $y \in Y$ ist $f^{-1}(\{y\}) \in \mathcal{A}$.*
- *$\mu(f^{-1}(Y \backslash \{\mathbf{0}\})) < \infty$.*

$$\mathcal{EF}(X, \mu, Y) := \{f : X \to Y : f \text{ ist } \mu\text{-einfach } \}$$

*bezeichne die Menge aller $\mu$-einfachen Funktionen. $f : X \to Y$ heißt* $\mu$**-meßbar**, *wenn es eine Folge $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{EF}(X, \mu, Y)$ gibt, so daß $f(x) = \lim\limits_{n\to\infty} f_n(x)$ für $\mu$-fast alle $x \in X$ gilt.*

$$\mathcal{L}_0(X, \mu, Y) := \{f : X \to Y : f \text{ ist } \mu\text{-meßbar}\}$$

*$f : X \to Y$ heißt* $\mu$**-fast separabelwertig**, *wenn es eine $\mu$-Nullmenge $N \in \mathcal{A}$ gibt, so daß $f(X \backslash N) \subseteq Y$ separabel ist (dh. daß es eine abzählbare, dichte Teilmenge von $f(X \backslash N)$ gibt.)*

**Satz 14.5.2.** *Es sei $(X, \mathcal{A}, \mu)$ ein vollständiger, $\sigma$-endlicher Maßraum, $Y$ ein Banachraum und $f : X \to Y$ eine Abbildung, dann sind äquivalent:*

*a)* $f : X \to Y$ *ist $\mu$-meßbar*

*b)* $f : X \to Y$ *ist $\mathcal{A} - \mathcal{B}(Y)$-meßbar und $\mu$-fast separabelwertig.*

**Korollar 14.5.3.** *Es sei $(X, \mathcal{A}, \mu)$ ein vollständiger, $\sigma$-endlicher Maßraum, $Y$ ein separabler Banachraum, dann sind äquivalent:*

*a)* $f : X \to Y$ *ist $\mu$-meßbar.*

*b)* $f : X \to Y$ *ist $\mathcal{A} - \mathcal{B}(Y)$-meßbar.*

Wie beim Beweis von Lemma 14.3.1 zeigt man auch für $\mu-$einfache Funktionen, daß die folgende Definition eines Integrals nicht von der Schreibweise der einfachen Funktion als Linearkombination von charakteristischen Funktionen abhängt:

**Definition 14.5.4.** *Es sei $(X, \mathcal{A}, \mu)$ ein vollständiger, $\sigma-$endlicher Maßraum, $Y$ ein Banachraum und $f = \sum\limits_{k=1}^{n} y_k \mathbf{1}_{A_k} \in \mathcal{EF}(X, \mu, Y)$ mit $y_1, ..., y_n \in Y$ und $A_1, ..., A_n \in \mathcal{A}$ eine $\mu-$einfache Funktion, dann heißt*

$$\int\limits_X f d\mu := \sum_{k=1}^{n} y_k \mu(A_k) \tag{14.5.1}$$

*das* **Integral** *von $f$.*

**Lemma 14.5.5.** *Es sei $(X, \mathcal{A}, \mu)$ ein vollständiger, $\sigma-$endlicher Maßraum, $Y$ ein Banachraum, dann gilt:*

*a)* $$\begin{array}{rcll} \int\limits_X \cdot\, d\mu : \mathcal{EF}(X,\mu,Y) & \to & Y & \quad \textit{ist linear.} \\ f & \mapsto & \int\limits_X f d\mu & \end{array}$$

*b) Für $f \in \mathcal{EF}(X,\mu,Y)$ ist*

$$\left\| \int\limits_X f d\mu \right\| \leq \int\limits_X \|f\| d\mu \tag{14.5.2}$$

*Beweis.* Der Beweis von a) geht wie bei Lemma 14.3.3; aus der Dreiecksungleichung in $Y$ folgt b). $\square$

**Bemerkung 14.5.6.**

- Auf $\mathcal{EF}(X,\mu,Y)$ wird durch

$$\|\phi\|_{\mathcal{EF}} := \int\limits_X \|\phi(x)\| d\mu(x)$$

eine Halbnorm definiert, aber wenn es in $\mathcal{A}$ eine $\mu-$Nullmenge $N \neq \emptyset$ gibt, dann ist $\|\cdot\|_{\mathcal{EF}}$ schon keine Norm mehr, denn $\|\mathbf{1}_N\|_{\mathcal{EF}} = 0$, aber $\mathbf{1}_N \neq 0$.

- Auch eine Halbnorm $\|\cdot\| : V \to [0, \infty[$ auf einem Vektorraum macht $V$ durch

$$\begin{array}{rcl} \mathcal{O}_V & := & \{U \subseteq V : \text{ Für jedes } x \in U \text{ gibt es } r = r(x) > 0 \text{ mit} \\ & & \qquad \{v \in V : \|x - v\| < r\} \subseteq U\} \end{array}$$

zu einem topologischen Raum. Ist aber $\|\cdot\|$ keine Norm, dann ist $(V, \mathcal{O}_V)$ iA. nicht hausdorffsch, zB. wenn $(X, \mathcal{A}, \mu)$ ein Maßraum, $x \in X$, $\{x\} \in \mathcal{A}$ mit $\mu(\{x\}) = 0$ ist, dann ist $0 \neq \mathbf{1}_{\{x\}}$ aber $\|0 - \mathbf{1}_{\{x\}}\|_{\mathcal{EF}} = 0$, dh. in jeder Umgebung von 0 bzgl. $\mathcal{O}_{\mathcal{EF}}$ ist auch $\mathbf{1}_{\{x\}}$ enthalten. Insbesondere brauchen Grenzwerte von **Cauchyfolgen** – und die sind genauso definiert wie bei einer Norm – bei einer Halbnorm nicht eindeutig sein.

**Lemma 14.5.7.** *Es sei $(X, \mathcal{A}, \mu)$ ein vollständiger, $\sigma-$endlicher Maßraum und $Y$ ein Banachraum. Sind $(\varphi_n)_{n\in\mathbb{N}}$ und $(\psi_n)_{n\in\mathbb{N}}$ zwei $\|\cdot\|_{\mathcal{EF}}-$Cauchyfolgen in $\mathcal{EF}(X,\mu,Y)$ mit $\lim\limits_{n\to\infty} \varphi_n(x) = \lim\limits_{n\to\infty} \psi_n(x)$ für $\mu-$fast alle $x \in X$, dann konvergieren die Folgen $\left(\int\limits_X \varphi_n d\mu\right)_{n\in\mathbb{N}}$ und $\left(\int\limits_X \psi_n d\mu\right)_{n\in\mathbb{N}}$ in $Y$ und es gilt:*

$$\lim_{n\to\infty} \int\limits_X \varphi_n d\mu = \lim_{n\to\infty} \int\limits_X \psi_n d\mu. \tag{14.5.3}$$

**Definition 14.5.8.** *Es sei $(X, \mathcal{A}, \mu)$ ein vollständiger, $\sigma-$endlicher Maßraum, $Y$ ein Banachraum, dann heißt $f : X \to Y$ $\mu-$***(Bochner-)integrierbar***, wenn es eine $\|\cdot\|_{\mathcal{EF}}$-Cauchyfolge $(\varphi_n)_{n\in\mathbb{N}}$ in $\mathcal{EF}(X, \mu, Y)$ gibt mit $f(x) = \lim\limits_{n\to\infty} \varphi_n(x)$ für $\mu-$fast alle $x \in X$. In diesem Fall ist nach Lemma 14.5.7*

$$\int\limits_X f d\mu := \lim_{n\to\infty} \int\limits_X \varphi_n d\mu \in Y \tag{14.5.4}$$

*wohldefiniert und heißt das* **(Bochner-$\mu-$)Integral** *von $f$. Schreibweise:*

$$\mathcal{L}^1(X, \mu, Y) := \{f : X \to Y : f \text{ ist } \mu-\text{integrierbar}\} \tag{14.5.5}$$

**Satz 14.5.9.** *Es sei $(X, \mathcal{A}, \mu)$ ein vollständiger, $\sigma-$endlicher Maßraum, $Y$ ein Banachraum und $f : X \to Y$ sei $\mu-$meßbar. Dann sind äquivalent:*

*a) $f$ ist (Bochner-$\mu$-)integrierbar*

*b) $\|f\| \in \mathcal{L}^1(X, \mu, [0, \infty])$.*

**Bemerkung 14.5.10.**

- $\mathcal{L}^1(X, \mu, Y)$ ist ein Vektorraum, denn zu $f, g \in \mathcal{L}^1(X, \mu, Y)$ und $\alpha \in \mathbb{K}$ gibt es $\|\cdot\|_{\mathcal{EF}}-$Cauchyfolgen $(\varphi_n)_{n\in\mathbb{N}}$ und $(\Psi_n)_{n\in\mathbb{N}}$ in $\mathcal{EF}(X, \mu, Y)$ mit $f(x) = \lim\limits_{n\to\infty} \varphi_n(x)$ und $g(x) = \lim\limits_{n\to\infty} \psi_n(x)$ für $\mu-$fast alle $x \in X$. Dann ist auch $(\varphi_n + \alpha\psi_n)_{n\in\mathbb{N}}$ eine $\|\cdot\|_{\mathcal{EF}}-$Cauchyfolge in $\mathcal{EF}(X, \mu, Y)$ mit $f(x) + \alpha g(x) = \lim\limits_{n\to\infty} (\varphi_n(x) + \alpha\psi_n(x))$ für $\mu-$fast alle $x \in X$ und damit ist $f + \alpha g \in \mathcal{L}^1(X, \mu, Y)$.

- Auf $\mathcal{L}^1(X, \mu, Y)$ definiert

$$\|f\|_{\mathcal{L}^1} := \int\limits_X \|f\| d\mu \tag{14.5.6}$$

  eine Halbnorm, denn zu $f, g \in \mathcal{L}^1(X, \mu, Y)$ wähle $\|\cdot\|_{\mathcal{EF}}-$Cauchyfolgen $(\varphi_n)_{n\in\mathbb{N}}$ und $(\psi_n)_{n\in\mathbb{N}}$ in $\mathcal{EF}(X, \mu, Y)$ mit $f(x) = \lim\limits_{n\to\infty} \psi_n(x)$ und $g(x) = \lim\limits_{n\to\infty} \varphi_n(x)$ für $\mu-$fast alle $x \in X$, dann gilt:

  - $\|f\|_{\mathcal{L}^1} = \int\limits_X \|f(x)\| d\mu(x) = \lim\limits_{n\to\infty} \int\limits_X \|\psi_n\| d\mu = \lim\limits_{n\to\infty} \|\psi_n\|_{\mathcal{EF}} \geq 0.$
  - Für $\alpha \in \mathbb{K}$ ist $\|\alpha f\|_{\mathcal{L}^1} = \lim\limits_{n\to\infty} \|\alpha\psi_n\|_{\mathcal{EF}} = |\alpha| \lim\limits_{n\to\infty} \|\psi_n\|_{\mathcal{EF}} = |\alpha|\|f\|_{\mathcal{L}^1}$
  - $\|f + g\|_{\mathcal{L}^1} = \lim\limits_{n\to\infty} \|\psi_n + \varphi_n\|_{\mathcal{EF}} \leq \lim\limits_{n\to\infty} (\|\psi_n\|_{\mathcal{EF}} + \|\varphi_n\|_{\mathcal{EF}}) = \|f\|_{\mathcal{L}^1} + \|g\|_{\mathcal{L}^1}$

**Satz 14.5.11.** *Es sei $(X, \mathcal{A}, \mu)$ ein vollständiger, $\sigma-$endlicher Maßraum und $Y$ ein Banachraum, dann gilt:*

*a)*
$$\begin{aligned} \int_X \cdot \, d\mu : \mathcal{L}^1(X,\mu,Y) &\to Y \\ f &\mapsto \int_X f d\mu \end{aligned}$$
*ist linear und stetig und es gilt:*

$$\left\| \int_X f d\mu \right\| \leq \int_X \|f\| d\mu = \|f\|_{\mathcal{L}^1} \tag{14.5.7}$$

*b) Ist $Z$ ein Banachraum, $T \in L(Y,Z)$, dann ist für jedes $f \in \mathcal{L}^1(X,\mu,Y)$ auch $Tf \in \mathcal{L}^1(X,\mu,Z)$ und es gilt:*

$$T\left[\int_X f d\mu\right] = \int_X T[f(x)] d\mu(x). \tag{14.5.8}$$

**Korollar 14.5.12.** *Es sei $(X,\mathcal{A},\mu)$ ein vollständiger, $\sigma-$endlicher Maßraum, $d \in \mathbb{N}$ und $f : X \to \mathbb{K}^d$ sei meßbar. $f = (f_1,...,f_d) : X \to \mathbb{K}^d$ ist genau dann $\mu-$integrierbar, wenn jede Koordinatenfunktion $f_j : X \to \mathbb{K}$, $j = 1,...,d$ $\mu-$integrierbar ist. In diesem Fall ist*

$$\int_X f d\mu = \left( \int_X f_1 d\mu, ... \int_X f_d d\mu \right) \tag{14.5.9}$$

**Satz 14.5.13** (majorisierte Konvergenz für banachraumwertige Funktionen)**.** *Es sei $(X,\mathcal{A},\mu)$ ein vollständiger, $\sigma-$endlicher Maßraum und $Y$ ein Banachraum. $(f_n)_{n\in\mathbb{N}}$ sei eine Folge in $\mathcal{L}^1(X,\mu,Y)$ und $g : X \to [0,\infty]$ sei $\mu-$integrierbar mit $\|f_n(x)\| \leq g(x)$ für $\mu-$fast alle $x \in X$. Der punktweise Grenzwert $f(x) := \lim\limits_{n\to\infty} f_n(x)$ existiere für alle $x \in X$, dann gilt:*

- $\begin{aligned} f : X &\to Y \\ x &\mapsto f(x) = \lim\limits_{n\to\infty} f_n(x) \end{aligned}$ *ist $\mu-$integrierbar*
- $\lim\limits_{n\to\infty} \|f_n - f\|_{\mathcal{L}^1} = 0$
- $\int_X f d\mu = \lim\limits_{n\to\infty} \int_X f_n d\mu.$

# Kapitel 15

# Differentialrechnung

## 15.1 Die Ableitung einer stetigen Funktion

**Definition 15.1.1.** *Es seien $X$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$ $(= \mathbb{R}$ oder $\mathbb{C})$, $U \subseteq X$ offen und $a \in U$. Zwei Funktionen $f : U \to Y$ und $g : U \to Y$* **berühren sich im Punkt** *$a$, wenn*

$$\lim_{\substack{x\to a\\ x\neq a}} \frac{\|f(x) - g(x)\|}{\|x - a\|} = 0 \tag{15.1.1}$$

*ist.*

**Bemerkung 15.1.2.** Es seien $X$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen und $a \in U$.

a) Berühren sich zwei in $a$ stetige Funktionen $f : U \to Y$ und $g : U \to Y$ in einem Punkt $a \in U$, so ist $f(a) = g(a)$, denn für die in $a$ stetige Funktion $f : U \to Y$ existiert $\lim\limits_{x\to a} f(x)$ und es gilt $f(a) = \lim\limits_{x\to a} f(x)$. Ist nun $f(a) \neq g(a)$, so ist $\lim\limits_{x\to a}(f(x) - g(x)) \neq 0$ und daher kann $\lim\limits_{\substack{x\to a\\ x\neq a}} \frac{\|f(x) - g(x)\|}{\|x - a\|}$ nicht existieren.

b) Berühren sich $f : U \to Y$ und $g : U \to Y$ im Punkt $a \in U$ und berühren sich $g$ und $h : U \to Y$ im Punkt $a$, so berühren sich wegen

$$\|f(x) - h(x)\| \leq \|f(x) - g(x)\| + \|g(x) - h(x)\|$$

auch die Funktionen $f$ und $h$ im Punkt $a$.

**Lemma 15.1.3.** *Es seien $X$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen, $a \in U$ und $f : U \to Y$ stetig in $a$. Dann gibt es in der Menge*

$$\{g : U \to Y : g \text{ berührt } f \text{ im Punkt } a\}$$

*höchstens eine Abbildung der Form* $\begin{array}{rcl} g : U & \to & Y \\ x & \mapsto & f(a) + u[x-a] \end{array}$*, wobei $u : X \to Y$ eine $\mathbb{K}$-lineare Abbildung ist.* [1]

[1] Wir werden im Folgenden oft Familien von linearen Abbildungen betrachten und schreiben wie hier das Argument, in dem die Funktionen linear sind in eckige Klammern

*Beweis.* Ist $w : X \to Y$ eine $\mathbb{K}$-lineare Abbildung, so daß $\begin{array}{rcl} h : U & \to & Y \\ x & \mapsto & f(a) + w[x-a] \end{array}$ im Punkt $a$ die Funktion $f$ berührt, so berühren sich auch $g$ und $h$ im Punkt $a$, dh.

$$0 = \lim_{\substack{x \to a \\ x \neq a}} \frac{\|u[x-a] - w[x-a]\|}{\|x-a\|} = \lim_{\substack{v \to 0 \\ v \neq \mathbf{0}}} \frac{\|(u-w)[v]\|}{\|v\|},$$

deshalb gibt es für jedes $\varepsilon > 0$ ein $r(\varepsilon) > 0$ mit $\|(u-w)[v]\| < \varepsilon\|v\|$ für alle $v \in X \backslash \{\mathbf{0}\}$ mit $\|v\| \leq r(\varepsilon)$. Wendet man diese Abschätzung auf $v := \frac{r(\varepsilon)x}{\|x\|}$ an, so folgt

$$\|(u-w)[x]\| < \varepsilon\|x\| \tag{15.1.2}$$

für alle $x \in X \backslash \{\mathbf{0}\}$. Da wir (15.1.2) für jedes $\varepsilon > 0$ gezeigt haben, folgt $\|(u-w)[x]\| = 0$ für jedes $x \in X \backslash \{\mathbf{0}\}$. Da $u - w$ linear ist, gilt $(u-w)[\mathbf{0}] = \mathbf{0}$, daher ist $u = w$. □

**Definition 15.1.4.** *Es seien $X$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen und $a \in U$. Eine in $a$ stetige Funktion $f : U \to Y$ heißt* **differenzierbar im Punkt** *$a$, wenn eine lineare Abbildung $u : X \to Y$ existiert, so daß die Abbildung*

$$\begin{array}{rcl} g : U & \to & Y \\ x & \mapsto & f(a) + u[x-a] \end{array} \tag{15.1.3}$$

*die Abbildung $f$ im Punkt $a$ berührt. Ist $f$ in $a$ differenzierbar, so ist die lineare Abbildung $u$ aus (15.1.3) nach Lemma 15.1.3 eindeutig bestimmt und heißt die* **(totale) Ableitung** *von $f$ im Punkt $a$. Schreibweise: $f'(a)$ oder $(Df)(a)$ statt $u$.*

**Lemma 15.1.5.** *Sind $X$ und $Y$ $\mathbb{K}$-Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen und $f : U \to Y$ und $g : U \to Y$ differenzierbar in $a \in U$. Dann ist für jedes $\lambda, \mu \in \mathbb{K}$ die Abbildung $\begin{array}{rcl} \lambda f + \mu g : U & \to & Y \\ x & \mapsto & \lambda f(x) + \mu g(x) \end{array}$ in $a$ differenzierbar mit*

$$(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a). \tag{15.1.4}$$

*Beweis.* Aufgabe Tutorium □

**Satz 15.1.6.** *Es seien $X$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen und die in $a$ stetige Funktion $f : U \to Y$ sei in $a \in U$ differenzierbar. Dann ist $f'(a) \in L(X,Y)$ eine stetige lineare Abbildung.*

*Beweis.* Da $f$ in $a$ stetig und in $a$ differenzierbar ist, gibt es für jedes $\varepsilon > 0$ ein $r(\varepsilon) \in ]0,1[$, so daß

$$\begin{array}{rcll} \|f(a+x) - f(a)\| & \leq & \dfrac{\varepsilon}{2} & \text{für alle } x \in X \text{ mit } \|x\| < r(\varepsilon) \\ \|f(a+x) - f(a) - (f'(a))[x]\| & \leq & \dfrac{\varepsilon}{2}\|x\| & \text{für alle } x \in X\backslash\{\mathbf{0}\} \text{ mit } \|x\| < r(\varepsilon). \end{array}$$

Deshalb ist

$$\|(f'(a))[x]\| = \Big\| f(a+x) - f(a) - \Big(f(a+x) - f(a) - (f'(a))[x]\Big)\Big\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2}\|x\| < \varepsilon$$

für alle $\|x\| < r(\varepsilon)$, dh. $f'(a)$ ist stetig bei $\mathbf{0}$ und daher $f'(a) \in L(X,Y)$ nach 13.9.1. □

**Lemma 15.1.7.** *Es seien $X$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen, $a \in U$ und $f : U \to Y$. Dann sind äquivalent:*

*a) $f : U \to Y$ ist in $a \in U$ differenzierbar.*

*b) Es gibt ein $T \in L(X,Y)$ und $r_a : U \to Y$, das in $a$ stetig ist, mit $r_a(a) = \mathbf{0}$ und es gilt*

$$f(x) = f(a) + T[x-a] + r_a(x)\,\|x-a\| \tag{15.1.5}$$

*für alle $x \in U$.*

*In diesem Fall ist $T = (Df)(a)$.*

*Beweis.*

a)⇒b) Ist $f$ in $a \in U$ differenzierbar, dann ist

$$\begin{aligned} r_a : U &\to Y \\ x &\mapsto \begin{cases} \dfrac{f(x) - f(a) - (Df)(a)[x-a]}{\|x-a\|} & \text{falls } x \neq a \\ \mathbf{0} & \text{falls } x = a \end{cases} \end{aligned}$$

in $a$ stetig, denn wegen der Differenzierbarkeit von $f$ in $a$ gilt $\lim\limits_{\substack{x \to a \\ x \neq a}} \|r_a(x)\| = 0$. Nach Definition von $r_a$ gilt ferner

$$f(x) = f(a) + (Df)(a)[x-a] + r_a(x)\,\|x-a\|$$

für jedes $x \in U$.

b)⇒a) Dann ist

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{\|f(x) - f(a) - T[x-a]\|}{\|x-a\|} = \lim_{\substack{x \to a \\ x \neq a}} \|r_a(x)\| = \|r_a(a)\| = 0,$$

also $f$ in $a$ differenzierbar mit $T = (Df)(a)$. □

**Beispiel 15.1.8.**

a) Sind $X, Y$ $\mathbb{K}$-Banachräume, $U \subseteq X$ offen, dann ist eine konstante Funktion $\begin{aligned} f : U &\to Y \\ x &\mapsto c \end{aligned}$ in jedem Punkt $a \in U$ differenzierbar mit $f'(a) = 0$ – denn in Lemma 15.1.7 können wir $r_a = 0$ wählen und alles paßt.

b) Eine stetige lineare Abbildung $u \in L(X,Y)$ ist in jedem Punkt $a \in X$ differenzierbar und es gilt

$$(Du)(a) = u, \tag{15.1.6}$$

denn für die lineare Abbildung $u$ gilt: $u[a] + u[x-a] = u[x]$, deshalb folgt die Behauptung aus Lemma 15.1.7 – wieder mit $r_a = 0$.

**Satz 15.1.9.** *Es sei $U \subseteq \mathbb{K}$ offen, $X$ ein $\mathbb{K}$-Banachraum, $a \in U$ und $f : U \to X$ stetig in $a$. Dann sind äquivalent:*

*a) $f$ ist differenzierbar in $a$.*

*b) In $X$ existiert $\lim\limits_{\substack{x\to a\\ x\neq a}} \dfrac{f(x)-f(a)}{x-a}$.*

*In diesem Fall gilt – nach der Identifikation von $\tilde{\xi} : \begin{array}{ccc} \mathbb{K} & \to & X \\ x & \mapsto & x\xi \end{array} \in L(\mathbb{K}, X)$ mit $\xi \in X$:*

$$f'(a) = \lim_{\substack{x\to a\\ x\neq a}} \frac{f(x)-f(a)}{x-a} \in X \tag{15.1.7}$$

*Beweis.*

a) $\Rightarrow$ b) Ist $f$ in $a$ differenzierbar, so gibt es genau ein $\tilde{u} \in L(\mathbb{K}, X)$ mit

$$\lim_{\substack{x\to a\\ x\neq a}} \frac{\|f(x)-f(a)-\tilde{u}[x-a]\|}{|x-a|} = 0.$$

Da $\tilde{u} \in L(\mathbb{K}, X)$ von der Form $\tilde{u} : \begin{array}{ccc} \mathbb{K} & \to & X \\ x & \mapsto & xu \end{array}$ mit $u = \tilde{u}(1) \in X$ ist, gilt:

$$\frac{\|f(x)-f(a)-\tilde{u}[x-a]\|}{|x-a|} = \left\| \frac{f(x)-f(a)}{x-a} - u \right\| \xrightarrow{\substack{x\to a\\ x\neq a}} 0,$$

dh. $\lim\limits_{\substack{x\to a\\ x\neq a}} \dfrac{f(x)-f(a)}{x-a} = u$ existiert in $X$.

b) $\Rightarrow$ a) Existiert $u := \lim\limits_{\substack{x\to a\\ x\neq a}} \dfrac{f(x)-f(a)}{x-a}$ in $X$, so definiert $\tilde{u} : \begin{array}{ccc} \mathbb{K} & \to & X \\ x & \mapsto & xu \end{array}$ eine stetige lineare Abbildung und

$$\frac{\|f(x)-f(a)-\tilde{u}[x-a]\|}{|x-a|} = \left\| \frac{f(x)-f(a)}{x-a} - u \right\| \xrightarrow{\substack{x\to a\\ x\neq a}} 0,$$

dh. $g : \begin{array}{ccc} U & \to & X \\ x & \mapsto & f(a)+\tilde{u}[x-a] \end{array}$ berührt $f$ im Punkt $a \in U$, dh. $f$ ist differenzierbar in $a$.

$\square$

**Beispiel 15.1.10.** a) Für jedes $n \in \mathbb{N}$ ist die Funktion $F_n : \begin{array}{ccc} \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & z^n \end{array}$ in jedem Punkt $a \in \mathbb{C}$ differenzierbar, denn

$$\frac{z^n - a^n}{z-a} = z^{n-1} + z^{n-2}a + ... + za^{n-2} + a^{n-1} \xrightarrow{z\to a} na^{n-1} = F'(a).$$

b) Die Funktion $\begin{array}{rcl} f:\mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \begin{cases} 0 & \text{falls } x=0 \\ x^2\sin(\frac{1}{x}) & \text{falls } x\neq 0 \end{cases} \end{array}$ ist in $a=0$ differenzierbar mit $f'(0)=0$, denn wegen $|\sin(y)|\leq 1$ für alle $y\in\mathbb{R}$ gilt:

$$\frac{f(x)-f(0)}{x} = x\sin(\frac{1}{x}) \overset{x\to 0, x\neq 0}{\longrightarrow} 0$$

c) Die Funktion $\begin{array}{rcl} f:\mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \begin{cases} 0 & \text{falls } x=0 \\ x\sin(\frac{1}{x}) & \text{falls } x\neq 0 \end{cases} \end{array}$ ist in $a=0$ nicht differenzierbar, denn $\dfrac{f(x)-f(0)}{x-0} = \sin(\dfrac{1}{x})$ hat in $a=0$ wegen $0=\sin(k\pi)=\sin\left(\frac{1}{\frac{1}{k\pi}}\right)$ und $1=\sin((4k+1)\frac{\pi}{2})=\sin\left(\frac{1}{\frac{2}{(4k+1)\pi}}\right)$ für $k\in\mathbb{N}$ keinen Grenzwert.

**Satz 15.1.11** (Kettenregel)**.** *Es seien $X, Y$ und $Z$ Banachräume über demselben Körper $\mathbb{K}$, $a\in X$, $U\subseteq X$ offen mit $a\in U$, $f:U\to Y$ stetig in $a$, $b:=f(a)$ und $V\subseteq Y$ offen mit $b\in V$, ferner $g:V\to Z$ stetig in $b$. Ist $f$ differenzierbar in $a$ und $g$ differenzierbar in $b=f(a)$, $f(U)\subseteq V$, so ist auch $h=g\circ f$ differenzierbar in $a$ und es gilt:*

$$(g\circ f)'(a) = g'(f(a))\circ f'(a). \tag{15.1.8}$$

*Beweis.* Nach Voraussetzung gibt es nach Lemma 15.1.7 für jedes $\varepsilon\in]0,1[$ ein $r=r(\varepsilon)\in ]0,1[$, so daß für

$$\begin{array}{rcl} \mathcal{R}_1(x) & := & f(a+x)-f(a)-f'(a)[x] \\ \mathcal{R}_2(y) & := & g(b+y)-g(b)-g'(b)[y] \end{array}$$

und für $\|x\|,\|y\|\leq r(\varepsilon)$ gilt:

$$\|\mathcal{R}_1(x)\|\leq\varepsilon\|x\| \qquad \text{und} \qquad \|\mathcal{R}_2(y)\|\leq\varepsilon\|y\|. \tag{15.1.9}$$

Wegen $f'(a)\in L(X,Y)$ und $g'(b)\in L(Y,Z)$ gilt für $C:=\max\{|||f'(a)|||,|||g'(b)|||\}$, daß

$$\|f'(a)[x]\|\leq C\|x\| \qquad \text{und} \qquad \|g'(b)[y]\|\leq C\|y\|. \tag{15.1.10}$$

Für $\|x\|\leq\dfrac{r(\varepsilon)}{1+C}$ ist

$$\begin{array}{rcl} g(f(a+x)) & = & g(f(a)+f'(a)[x]+\mathcal{R}_1(x)) \\ & = & g(f(a))+g'(f(a))[f'(a)[x]+\mathcal{R}_1(x)]+\mathcal{R}_2(f'(a)[x]+\mathcal{R}_1(x)) \\ & = & g(b)+g'(b)[f'(a)[x]]+g'(b)[\mathcal{R}_1(x)]+\mathcal{R}_2(f'(a)[x]+\mathcal{R}_1(x)). \end{array}$$

Wegen $\|x\|\leq\dfrac{r(\varepsilon)}{1+C}$ ist $\|f'(a)[x]+\mathcal{R}_1(x)\|\leq C\|x\|+\varepsilon\|x\|\leq(C+1)\|x\|\leq r(\varepsilon)$ und damit folgt $\|\mathcal{R}_2(f'(a)[x]+\mathcal{R}_1(x))\|\leq\varepsilon\|f'(a)[x]+\mathcal{R}_1(x)\|\leq\varepsilon(C+1)\|x\|$ aus (15.1.9). Ferner ist $\|g'(b)[\mathcal{R}_1(x)]\|\leq C\|\mathcal{R}_1(x)\|\leq C\varepsilon\|x\|$, daher ist

$$\begin{array}{l} \left\|g(f(a+x))-g(f(a))-g'(f(a))\left[f'(a)[x]\right]\right\| \\ \quad\leq\ \|(g'(b))[\mathcal{R}_1(x)]+\mathcal{R}_2(f'(a)[x]+\mathcal{R}_1(x))\|\leq\varepsilon(2C+1)\|x\| \end{array} \tag{15.1.11}$$

für alle $\|x\| \leq \dfrac{r(\varepsilon)}{1+C}$. Damit ist

$$\lim_{\substack{x\to 0\\ x\neq 0}} \frac{\|g(f(a+x)) - g(f(a)) - (g'(f(a)) \circ f'(a))[x]\|}{\|x-a\|} = 0,$$

also $g \circ f$ in $a$ differenzierbar mit $(g \circ f)'(a) = g'(f(a)) \circ f'(a)$. $\square$

**Satz 15.1.12.** *Es seien $X, Y = Y_1 \times ... \times Y_m$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen, $a \in U$ und für $l \in \{1, ..., m\}$ sei* $\begin{array}{rcl} \mathrm{pr}_l : Y & \to & Y_l \\ y = (y_1, ..., y_m) & \mapsto & y_l \end{array}$ *die kanonische Projektion von $Y$ auf den $l$-ten Faktor $Y_l$. Es sei $f : U \to Y$ stetig in $a$, dann sind äquivalent:*

*a) $f$ ist in $a \in U$ differenzierbar.*

*b) Für alle $l = 1, ..., m$ sind die Abbildungen* $\begin{array}{rcl} f_l := \mathrm{pr}_l \circ f : U & \to & Y_l \\ x & \mapsto & \mathrm{pr}_l(f(x)) \end{array}$ *in $a \in U$ differenzierbar.*

*In diesem Fall ist $f'(a) = (f_1'(a), ..., f_m'(a))$, wenn man $L(X, Y)$ mit $\prod_{j=1}^{m} L(X, Y_j)$ identifiziert.*

*Beweis.* Es seien $\begin{array}{rcl} \mathrm{pr}_l : Y_1 \times ... \times Y_m & \to & Y_l \\ y = (y_1, ..., y_m) & \mapsto & y_l \end{array}$ für $l = 1, ..., m$ die kanonischen Projektionen. Ist $u \in L(X, Y)$, so ist auch $u_l := \mathrm{pr}_l \circ u$ linear und nach Definition der Summennorm als $\|y\| = \|y_1\| + ... + \|y_m\|$ auf $Y$ folgt $\|u_l(x)\| \leq \|u(x)\| \leq |||u|||\,\|x\|$ für jedes $x \in X$ und daher ist $u_l \in L(X, Y_l)$. Sind umgekehrt $u_1 \in L(X, Y_1)$,..., $u_m \in L(X, Y_m)$, so ist $\begin{array}{rcl} u : X & \to & Y = Y_1 \times ... \times Y_m \\ x & \mapsto & u(x) := (u_1(x), ..., u_m(x)) \end{array}$ offenbar $\mathbb{K}$-linear und

$$\|u(x)\| = \|u_1(x)\| + ... + \|u_m(x)\| \leq (|||u_1||| + ... + |||u_m|||)\|x\|,$$

also $u \in L(X, Y)$. Das erlaubt es uns durch die bijektive lineare Abbildung

$$\begin{array}{rcl} L(X, Y) & \to & \prod_{l=1}^{m} L(X, Y_l) \\ u & \mapsto & (\mathrm{pr}_1 \circ u, ..., \mathrm{pr}_m \circ u) \end{array}$$

die beiden Räume $L(X, Y)$ und $\prod_{l=1}^{m} L(X, Y_l)$ miteinander zu identifizieren.

a)$\Rightarrow$b) Ist $f : U \to Y$ in $a \in U$ differenzierbar, so existiert nach Lemma 15.1.7 die Ableitung $(Df)(a) \in L(X, Y)$ und eine in $a$ stetige Funktion $r_a : U \to Y$ mit $r_a(a) = \mathbf{0}$ und

$$f(x) = f(a) + (Df)(a)[x-a] + r_a(x)\,\|x-a\|,$$

also ist

$$f_l(x) = (\mathrm{pr}_l \circ f)(a) + (\mathrm{pr}_l \circ (Df)(a))[x-a] + (\mathrm{pr}_l \circ r_a)(x)\,\|x-a\|$$

und daher $f_l$ in $a$ differenzierbar mit $(Df_l)(a) = (\mathrm{pr}_l \circ (Df)(a))$ nach Lemma 15.1.7, denn $\mathrm{pr}_l \circ r_a$ ist in $a$ stetig mit $(\mathrm{pr}_l \circ r_a)(a) = \mathbf{0}$.

b)⇒a) Sind die Abbildungen $f_l : U \to Y_l$ für jedes $l = 1, ..., m$ in $a$ differenzierbar, dann gibt es in $a$ stetige Funktionen $r_{a,l} : U \to Y_l$ mit $r_{a,l}(a) = \mathbf{0}$ und

$$f_l(x) = f_l(a) + (Df_l)(a)[x-a] + r_{a,l}(x)\,\|x-a\|.$$

Daraus folgt

$$\begin{array}{rcl} f(x) & = & (f_1(x), ..., f_m(x)) \\ & = & (f_1(a), ..., f_n(a)) + ((Df_1)(a)[x-a], ..., (Df_m)(a)[x-a]) \\ & & +(r_{a,1}(x), ..., r_{a,m}(x))\,\|x-a\| \end{array}$$

für jedes $x \in U$. Nach Definition der Produkttopologie ist

$$\begin{array}{rcl} r_a : U & \to & Y \\ x & \mapsto & (r_{a,1}(x), ..., r_{a,m}(x)) \end{array}$$

in $a$ stetig mit $r_a(a) = \mathbf{0}$, daher ist $((Df_1)(a), ..., (Df_m)(a)) \in \prod_{l=1}^{m} L(X, Y_l)$ nach Lemma 15.1.7 die Ableitung von $f$ in $a$. □

**Lemma 15.1.13.** *Es seien $X_1, ..., X_n, Y$ Banachräume über demselben Körper $\mathbb{K}$ und $\phi : X_1 \times ... \times X_n \to Y$ sei stetig und multilinear. Dann ist $\phi$ in jedem Punkt $a = (a_1, ..., a_n) \in X_1 \times ... \times X_n$ differenzierbar und*

$$\phi'(a_1, ..., a_n)[x_1, ..., x_n] = \sum_{j=1}^{n} \phi(a_1, ..., a_{j-1}, x_j, a_{j+1}, ..., a_n). \tag{15.1.12}$$

*Beweis.* Aufgabe □

**Korollar 15.1.14** (Produktregel)**.** *Es seien $X, Y$ zwei $\mathbb{K}$-Banachräume, $U \subseteq X$ offen, $a \in U$ und $f : U \to Y$, $g : U \to \mathbb{K}$ seien differenzierbar in $a$. Dann ist auch* $\begin{array}{rcl} f \cdot g : U & \to & Y \\ x & \mapsto & f(x)g(x) \end{array}$ *in $a$ differenzierbar mit $(f \cdot g)'(a) = f(a)g'(a) + g(a)f'(a)$.*

*Beweis.* Die Skalarmultiplikation $\begin{array}{rcl} m : \mathbb{K} \times Y & \to & Y \\ (\lambda, y) & \mapsto & \lambda y \end{array}$ ist offenbar bilinear. Da nach Voraussetzung $f$ und $g$ in $a \in U$ differenzierbar sind, ist

$$\begin{array}{rcl} F = (f, g) : U & \to & \mathbb{K} \times Y \\ x & \mapsto & (g(x), f(x)) \end{array}$$

nach Satz 15.1.12 in $a$ differenzierbar. Nach der Kettenregel in Satz 15.1.11 folgt $(f \cdot g)'(a) = (m \circ F)'(a) = m'(F(a)) \circ F'(a) = f(a)g'(a) + g(a)f'(a)$. □

**Lemma 15.1.15** (Quotientenregel)**.** *Es sei $X$ ein Banachraum, $U \subseteq X$ offen, $a \in U$, $f : U \to \mathbb{K}$ und $g : U \to \mathbb{K}$ seien stetig und in $a$ differenzierbar mit $g(a) \neq 0$. Dann ist*

$$\begin{array}{rcl} \frac{f}{g} : V & \to & \mathbb{K} \\ x & \mapsto & \frac{f(x)}{g(x)} \end{array}$$

*auf der offenen Menge $V := g^{-1}(\mathbb{K}\backslash\{0\})$ wohldefiniert und in $a$ differenzierbar mit*

$$\left(\frac{f}{g}\right)'(a) = \frac{g(a)f'(a) - f(a)g'(a)}{(g(a))^2} \tag{15.1.13}$$

*Beweis.* Aufgabe □

**Satz 15.1.16.** *Es seien $X, Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen, $V \subseteq Y$ offen und $f : U \to V$ ein Homöomorphismus und $g = f^{-1} : V \to U$ der inverse Homöomorphismus. $f$ sei in $a \in U$ differenzierbar und $f'(a) : X \to Y$ sei ein linearer Homöomorphismus. Dann ist $g$ im Punkt $b := f(a)$ differenzierbar und $g'(b)$ ist die zu $f'(a)$ inverse Abbildung, dh.*

$$\left((f^{-1})'\right)(f(a)) = (f'(a))^{-1}. \tag{15.1.14}$$

*Beweis.* Nach Voraussetzung gibt es die offenen Umgebungen $\tilde{U} = U - a$ von $\mathbf{0}$ in $X$ und $\tilde{V} = V - f(a)$ von $\mathbf{0}$ in $Y$, so daß $\begin{array}{rcl} \tilde{f} : \tilde{U} & \to & \tilde{V} \\ x & \mapsto & f(a+x) - f(a) \end{array}$ ein Homöomorphismus mit dem inversen [2] Homöomorphismus $\begin{array}{rcl} \tilde{g} : \tilde{V} & \to & \tilde{U} \\ y & \mapsto & g(b+y) - g(b) \end{array}$ ist. Nach Voraussetzung gibt es $u \in L(Y, X)$ mit $u \circ f'(a) = \mathrm{id}_X$, $f'(a) \circ u = \mathrm{id}_Y$. Da $f$ in $a$ differenzierbar ist, können wir zu jedem $\varepsilon \in ]0, \frac{1}{2|||u|||}[$ ein $r(\varepsilon) > 0$ wählen , so daß $\{x \in X : \|x\| < r(\varepsilon)\} \subseteq \tilde{U}$ und so daß für

$$\mathcal{R}(x) := f(a+x) - f(a) - f'(a)[x]$$

und alle $x \in X$ mit $\|x\| < r(\varepsilon)$ gilt:

$$\|\mathcal{R}(x)\| \leq \varepsilon \|x\|. \tag{15.1.15}$$

Da $\tilde{V}$ offen und $\tilde{g}$ stetig ist, gibt es $s(\varepsilon) > 0$, so daß $\{y \in Y : \|y\| < s(\varepsilon)\} \subseteq \tilde{V}$ und $\tilde{g}(\{y \in Y : \|y\| < s(\varepsilon)\}) \subseteq \{x \in X : \|x\| < r(\varepsilon)\}$. Für $\|y\| < s(\varepsilon)$ und $z := \tilde{g}(y)$ ist $y = \tilde{f}(z) = f(a+z) - f(a)$. Wegen $\|z\| < r(\varepsilon)$ ist also $y = f(a+z) - f(a) = f'(a)[z] + \mathcal{R}(z)$ mit $\|\mathcal{R}(z)\| \leq \varepsilon \|z\|$. Daher ist

$$u[y] = u[f'(a)[z] + \mathcal{R}(z)] = u \circ f'(a)[z] + u[\mathcal{R}(z)] = z + u[\mathcal{R}(z)],$$

daher ist wegen $\|u[\mathcal{R}(z)]\| \leq |||u||| \cdot \|\mathcal{R}(z)\| \leq \varepsilon |||u||| \|z\| < \frac{1}{2}\|z\|$

$$\|u[y]\| \geq \big|\|z\| - \|u[\mathcal{R}(z)]\|\big| \geq \frac{1}{2}\|z\|$$

oder

$$\|z\| \leq 2\|u[y]\| \leq 2|||u||| \|y\|.$$

Für $\|y\| < s(\varepsilon)$ ist also

$$\begin{aligned} \|g(b+y) - g(y) - u[y]\| &= \|\tilde{g}(y) - u[y]\| = \|z - (z + u[\mathcal{R}(z)])\| = \|u[\mathcal{R}(z)]\| \\ &\leq |||u||| \cdot \|\mathcal{R}(z)\| \leq \varepsilon |||u||| \cdot \|z\| = 2\varepsilon |||u|||^2 \|y\|. \end{aligned}$$

deshalb ist $g$ an der Stelle $b = f(a)$ differenzierbar mit $g'(b) = u$, dh es gilt (15.1.14). □

[2] denn für alle $x \in \tilde{U}$ und $y \in \tilde{V}$ gilt:

$$\begin{aligned} \tilde{f}(\tilde{g}(y)) = \tilde{f}(g(b+y) - g(b)) = f(a + f^{-1}(f(a)+y) - f^{-1}(f(a))) - f(a) = f(f^{-1}(f(a)+y)) - f(a) &= y \\ \tilde{g}(\tilde{f}(x)) = g(b + \tilde{f}(x)) - g(b) = f^{-1}(f(a) + f(a+x) - f(a)) - f^{-1}(f(a)) = f^{-1}(f(a+x)) - a &= x \end{aligned}$$

**Bemerkung 15.1.17.** In Satz 15.1.16 kann man auf die Voraussetzung „$f'(a)$ linearer Homöomorphismus“ nicht verzichten, zB. ist $\begin{array}{rcl} f:\mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^3 \end{array}$ stetig und bijektiv und die Umkehrabbildung $\begin{array}{rcl} g:\mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \sqrt[3]{x} \end{array}$ ist ebenfalls stetig. [3] Die Ableitung

$$\begin{array}{rcl} f'(a):\mathbb{R} & \to & \mathbb{R} \\ y & \mapsto & 3a^2y \end{array} \in L(\mathbb{R},\mathbb{R})$$

ist für $a = 0$ die Nullabbildung, also kein linearer Homöomorphismus und $g$ ist in 0 nicht differenzierbar, denn für $x > 0$ ist

$$\frac{g(x)-g(0)}{x-0} = x^{-\frac{2}{3}} \overset{x\searrow 0}{\longrightarrow} \infty,$$

also kann in $\mathbb{R}$ der Grenzwert $\lim\limits_{x\searrow 0} \dfrac{g(x)-g(0)}{x-0}$ nicht existieren, dh. nach Satz 15.1.9 ist $g$ in 0 nicht differenzierbar.

## 15.2 Hauptsatz der Differential- und Integralrechnung

**Satz 15.2.1** (Hauptsatz der Differential- und Integralrechnung)**.** *Es sei $Y$ ein Banachraum, es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \to Y$ sei stetig. Dann ist*

$$\begin{array}{rcl} F:[a,b] & \to & Y \\ x & \mapsto & \int\limits_a^x f(t)dt := \int\limits_{\mathbb{R}} \mathbf{1}_{[a,x]}(t)f(t)d\tilde{\lambda}(t) \end{array}$$

*wohldefiniert, $F$ ist stetig auf $[a, b]$ und stetig differenzierbar auf $]a, b[$ und für jedes $t \in ]a, b[$ gilt:*

$$F'(t) = f(t). \tag{15.2.1}$$

*Beweis.* Da $f : [a, b] \to Y$ stetig ist, ist $f([a, b])$ nach Lemma 13.6.3 kompakt, also insbesondere totalbeschränkt nach Satz 13.6.7. Für jedes $m \in \mathbb{N}$ gibt es also $M = M(m) \in \mathbb{N}$ und $y_1, ..., y_{M(m)} \in Y$ mit $f([a, b]) \subseteq \bigcup\limits_{k=1}^{M(m)} K(y_k, \dfrac{1}{m})$. Dann ist durch

$$\begin{array}{rcl} X_{1,m} & := & f^{-1}(K(y_1, \frac{1}{m})) \\ \vdots & & \vdots \\ X_{k,m} & := & f^{-1}(K(y_k, \frac{1}{m}))\setminus \left( \bigcup\limits_{j=1}^{k-1} X_{j,m} \right) \end{array}$$

[3] Wir hatten im Tutorium im 1. Semester bewiesen, daß für eine konvergente Folge $x_n \overset{n\to\infty}{\longrightarrow} x$ in $[0, \infty[$ auch $\sqrt{x_n} \overset{n\to\infty}{\longrightarrow} \sqrt{x}$ gilt, dh, daß die Quadratwurzel eine stetige Funktion ist; für die dritte Wurzel geht das analog.

eine Zerlegung von $[a,b] = \bigcup\limits_{k=1}^{M(m)} X_{k,m}$ in paarweise disjunkte meßbare Mengen gegeben. Daher definiert $f_m : [a,b] \to Y$ mit $f_m(x) = y_k$, falls $x \in X_{k,m}$ ist eine Folge $(f_m)_{m\in\mathbb{N}}$ von $\mu-$einfachen Funktionen mit $\|f_m(x) - f(x)\| \leq \frac{1}{m}$, dh. $f(x) = \lim\limits_{m\to\infty} f_m(x)$ für alle $x \in [a,b]$, dh. $f$ ist $\tilde{\lambda}-$meßbar. Da $f$ auf dem kompakten Intervall $[a,b]$ stetig ist, ist

$$\|f\|_\infty := \sup\{|f(t)| : t \in [a,b]\} < \infty$$

und damit $\|f\|_\infty \mathbf{1}_{[a,b]}$ eine integrierbare Majorante für alle Funktionen $\mathbf{1}_{[a,x]}f$, $x \in [a,b]$. Nach Satz 14.5.9 ist jedes $\mathbf{1}_{[a,x]}f$ $\tilde{\lambda}-$integrierbar und damit

$$\begin{array}{rcl} F : [a,b] & \to & Y \\ x & \mapsto & \int\limits_a^x f(t)dt := \int\limits_{\mathbb{R}} \mathbf{1}_{[a,x]}(t)f(t)d\tilde{\lambda}(t) \end{array}$$

wohldefiniert. Ist $(x_n)_{n\in\mathbb{N}}$ eine Folge in $[a,b]$ mit $\lim\limits_{n\to\infty} x_n = x \in [a,b]$, so gilt auch $\mathbf{1}_{[a,x]}(t)f(t) = \lim\limits_{n\to\infty} \mathbf{1}_{[a,x_n]}(t)f(t)$ für alle $t \in [a,b]$ und mit der $\tilde{\lambda}-$Majorante $\|f\|_\infty \mathbf{1}_{[a,b]}$ folgt aus dem Satz von der majorisierten Konvergenz:

$$F(x) = \int\limits_{[a,b]} \mathbf{1}_{[a,x]}(t)f(t)d\tilde{\lambda}(t) = \lim_{n\to\infty} \int\limits_{[a,b]} \mathbf{1}_{[a,x_n]}(t)f(t)d\tilde{\lambda}(t) = \lim_{n\to\infty} f(x_n),$$

dh. $F$ ist stetig. Weil $f$ auf dem kompakten Intervall $[a,b]$ als stetig vorausgesetzt wurde, ist $f$ nach Satz 13.6.10 sogar gleichmäßig stetig. Zu jedem $\varepsilon > 0$ gibt es ein $\delta = \delta(\varepsilon) > 0$, so daß für alle $x, y \in [a,b]$ mit $|x-y| < \delta$ gilt: $\|f(x) - f(y)\| < \varepsilon$. Ist nun $h \in \mathbb{R}$ mit $|h| < \delta$, dann ist für jedes $x \in ]a,b[$

$$\|F(x+h) - F(x) - hf(x)\| = \left\| \int\limits_x^{x+h} (f(t) - f(x))dt \right\| \leq \left| \int\limits_x^{x+h} \|f(t) - f(x)\|dt \right| \leq |h|\varepsilon,$$

was $F'(x) = f(x)$ im Limes $h \to 0$ ergibt. □

**Lemma 15.2.2.** *Es sei $Y$ ein Banachraum, $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a,b] \to Y$ stetig. Ist $G : [a,b] \to Y$ eine stetige Funktion mit $G'(x) = f(x)$ für jedes $x \in ]a,b[$, dann gilt*

$$\int\limits_c^d f(t)dt = G(d) - G(c) \tag{15.2.2}$$

*für jedes $c, d \in ]a,b[$.*

*Beweis.* Es sei

$$\begin{array}{rcl} F : [a,b] & \to & Y \\ x & \mapsto & \int\limits_a^x f(t)dt := \int\limits_{\mathbb{R}} \mathbf{1}_{[a,x]}(t)f(t)d\tilde{\lambda}(t) \end{array}$$

wie im Beweis von Satz 15.2.1. Ist nun $G : [a,b] \to Y$ eine Funktion mit $G'(t) = f(t)$ für jedes $t \in ]a,b[$, dann ist $G'(t) - F'(t) = (G-F)'(t) = 0$ für alle $t \in ]a,b[$ und weil $]a,b[$ zusammenhängend ist, ist $G - F$ eine konstante Funktion. Deshalb gilt nach Satz 15.2.1:

$$\int_c^d f(t)dt = F(d) - F(c) = G(d) - G(c)$$

für alle $c, d \in ]a,b[$. □

**Definition 15.2.3.** *Es sei $Y$ ein Banachraum, $f : [a,b] \to Y$ stetig, dann heißt eine Funktion $F : [a,b] \to Y$ mit $F'(t) = f(t)$ für $t \in ]a,b[$ eine* **Stammfunktion** *von $f$.*

**Satz 15.2.4** (Mittelwertsatz in Integralform)**.** *Es sei $Y$ ein Banachraum, $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a,b] \to Y$ stetig und $f|_{]a,b[}$ differenzierbar. Dann gilt*

$$f(b) - f(a) = \int_a^b f'(t)dt. \tag{15.2.3}$$

*Beweis.* $f$ ist eine Stammfunktion von $f'$. □

**Korollar 15.2.5** (partielle Integration)**.** *Es sei $Y$ ein $\mathbb{K}-$Banachraum, $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a,b] \to Y$, $g : [a,b] \to \mathbb{K}$ seien stetig und stetig differenzierbar auf $]a,b[$. Dann gilt für alle $c, d \in [a,b]$, $c < d$:*

$$\int_c^d f(t)g'(t)dt = f(d)g(d) - f(c)g(c) - \int_c^d f'(t)g(t)dt. \tag{15.2.4}$$

*Beweis.* Für $\begin{aligned} F : [a,b] &\to Y \\ t &\mapsto f(t)g(t) \end{aligned}$ berechnet sich die Ableitung nach der Produktregel als $F'(t) = f'(t)g(t) + f(t)g'(t)$ und damit folgt aus Satz 15.2.1:

$$\int_c^d f'(t)g(t)dt + \int_c^d f(t)g'(t)dt = F(d) - F(c).$$

□

**Korollar 15.2.6** (Substitutionsregel)**.** *Es sei $Y$ ein Banachraum, $a, b, c, d \in \mathbb{R}$ $a < b$ und $c < d$ und $f : [c,d] \to Y$ sei stetig und $\varphi : [a,b] \to [c,d]$ sei auf $]a,b[$ stetig differenzierbar. Dann gilt:*

$$\int_a^b f(\varphi(t))\varphi'(t)dt = \int_{\varphi(a)}^{\varphi(b)} f(x)dx. \tag{15.2.5}$$

*Beweis.* Ist $F : [c,d] \to Y$ eine Stammfunktion von $f$, so gilt für $F \circ \varphi : [a,b] \to Y$ nach der Kettenregel:

$$(F \circ \varphi)'(t) = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t)$$

und damit folgt nach Satz 15.2.1

$$\int\limits_a^b f(\varphi(t))\varphi'(t)dt = (F \circ \varphi)(b) - (F \circ \varphi)(a) = F(\varphi(b)) - F(\varphi(a)) = \int\limits_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

□

## 15.3 Schrankensatz

**Satz 15.3.1 (Mittelwertsatz in Integralform und Schrankensatz).** *Es seien $X, Y$ Banachräume, $U \subseteq X$ offen, $f : U \to Y$ stetig differenzierbar und $a, b \in U$, so daß die* **Verbindungsstrecke** $[\![a,b]\!] := \{a + t(b-a) : t \in [0,1]\}$ *in $U$ enthalten ist: $[\![a,b]\!] \subseteq U$, dann gilt:*

$$f(b) - f(a) = \int\limits_0^1 Df(a + t(b-a))[b-a]dt \tag{15.3.1}$$

*und insbesondere folgt:*

$$\|f(b) - f(a)\| \leq \|b - a\| \sup \Big\{|||f'(a + t(b-a))||| : t \in ]0,1[\Big\} \tag{15.3.2}$$

*Beweis.* Nach Voraussetzung ist

$$\begin{array}{rcl} \varphi : [0,1] & \to & Y \\ t & \mapsto & f(a + t(b-a)) \end{array}$$

stetig differenzierbar auf $]0,1[$ mit $\varphi'(t) = (Df)(a + t(b-a))[b-a]$ und stetig auf $[0,1]$. Nach dem Fundamentalsatz 15.2.1 der Differential- und Integralrechnung ist

$$f(b) - f(a) = \varphi(1) - \varphi(0) = \int\limits_0^1 \varphi'(t)dt = \int\limits_0^1 (Df)(a + t(b-a))[b-a]dt$$

und damit

$$\begin{array}{rcl} \|f(b) - f(a)\| & = & \left\| \int\limits_0^1 (Df)(a + t(b-a))[b-a]dt \right\| \\ & \leq & \|b - a\| \sup \{|||(Df)(a + t(b-a))||| : t \in ]0,1[\} \end{array}$$

□

**Lemma 15.3.2.** *Es seien $X, Y$ Banachräume über demselben Körper $\mathbb{K}$ und $U \subseteq X$ sei offen und zusammenhängend. $f : U \to Y$ sei stetig und in jedem Punkt $a \in U$ differenzierbar mit $f'(a) = \mathbf{0}$. Dann ist $f : U \to Y$ eine konstante Funktion.*

*Beweis.* Fixiere $a \in U$ und definiere

$$A := \{x \in U : f(x) = f(a)\} = U \cap f^{-1}(\{f(a)\})$$

dann ist $A$ bezüglich der Relativtopologie $\mathcal{O}_U$ von $U$ abgeschlossen nach Satz 13.1.15, denn $f$ ist stetig und $\{f(a)\}$ ist als einpunktige Teilmenge des Banachraums $Y$ abgeschlossen. Ist $x \in A \subseteq U$ und $r > 0$, so daß die offene Kugel $K(x, r) := \{y \in X : \|x - y\| < r\} \subseteq U$, dann gilt für jedes $y \in K(x, r)$ nach Satz 15.3.1

$$\|f(x) - f(y)\| \leq \|x - y\| \sup\{|||f'(x + t(y - x))||| : t \in [0, 1]\} = 0.$$

Mit jedem $x \in A$ ist also auch $K(x, r) \subseteq A$, daher ist $A$ in der Relativtopologie $\mathcal{O}_U$ auch offen. Da $U$ zusammenhängend ist, $a \in A \neq \emptyset$ und $A$ sowohl offen als auch abgeschlossen in $\mathcal{O}_U$ ist, folgt $A = U$ nach Lemma 13.5.2, dh. $f$ ist konstant auf $A = U$. □

**Bemerkung 15.3.3.** Ohne die Voraussetzung „$U$ zusammenhängend“ wird Lemma 15.3.2 falsch, zB. für $U = ]0, 1[ \cup ]1, 2[$ ist $\begin{array}{rcl} f : U & \to & \mathbb{R} \\ x & \mapsto & \begin{cases} 0 & \text{für } x \in ]0,1[ \\ 1 & \text{für } x \in ]1,2[ \end{cases} \end{array}$ in jedem Punkt $a \in U$ differenzierbar mit $f'(a) = 0$.

**Satz 15.3.4.** *Es seien $X, Y$ Banachräume über demselben Körper $\mathbb{K}$ und $U \subseteq X$ sei offen und zusammenhängend und $(f_n : U \to Y)_{n\in\mathbb{N}}$ eine Folge von differenzierbaren Abbildungen. Gilt dann noch*

*a) $(f'_n)_{n\in\mathbb{N}}$ konvergiert lokal gleichmäßig.*

*b) Es gibt ein $a \in U$, so daß die Folge $(f_n(a))_{n\in\mathbb{N}}$ in $Y$ konvergiert.*

*dann konvergiert die Folge $(f_n)_{n\in\mathbb{N}}$ lokal gleichmäßig auf $U$. Sind* $\begin{array}{rcl} f : U & \to & Y \\ x & \mapsto & \lim\limits_{n\to\infty} f_n(x) \end{array}$ *und* $\begin{array}{rcl} g : U & \to & L(X, Y) \\ x & \mapsto & \lim\limits_{n\to\infty} f'_n(x) \end{array}$*, dann gilt $g(x) = f'(x)$ für alle $x \in U$.*

*Beweis.* Wähle $a \in U$, so daß die Folge $(f_n(a))_{n\in\mathbb{N}}$ in $Y$ konvergiert und wähle $r > 0$, so daß $V := K(a, r) := \{x \in X : \|x - a\| < r\} \subseteq U$ ist und $(f'_n|_V)_{n\in\mathbb{N}}$ gleichmäßig konvergiert. Für jedes $x \in V$ liegt die verbindende Strecke $[\![a, x]\!]$ in $V$, daher gilt dann

$$\begin{aligned} \|f_n(x) - f_m(x) - (f_n(a) - f_m(a))\| &\leq \|x - a\| \sup\{|||f'_n(z) - f'_m(z)||| : z \in V\} \\ &\leq r \sup\{|||f'_n(z) - f'_m(z)||| : z \in V\} \end{aligned} \tag{15.3.3}$$

nach Satz 15.3.1. Weil $(f'_n|_V)_{n\in\mathbb{N}}$ gleichmäßig konvergiert, folgt die gleichmäßige Schranke

$$\sup\{|||f'_n(z) - f'_m(z)||| : z \in V\} \overset{m,n\to\infty}{\longrightarrow} 0$$

und da $(f_n(a))_{n\in\mathbb{N}}$ konvergiert folgt aus (15.3.3) gleichmäßige Konvergenz von $(f_n|_V)_{n\in\mathbb{N}}$. Es sei nun

$$A := \{x \in U : (f_n(x))_{n\in\mathbb{N}} \text{ konvergiert}\},$$

dann haben wir eben gezeigt, daß mit jedem $a \in A$ auch ein $r = r(a) > 0$ existiert, so daß $V := K(a, r(a)) \subseteq A$ gilt. Damit ist $A$ offen und nach Voraussetzung ist $a \in A \neq \emptyset$. Ist nun $x \in U$ ein Berührpunkt von $A$, so können wir weil $U$ offen ist, ein $s = s(x) > 0$ mit $K(x, s) \subseteq U$ wählen und oE. $s > 0$ so weit verkleinern, daß $(f'_n|_{K(x,s)})_{n\in\mathbb{N}}$ gleichmäßig konvergiert. Weil $K(x, s)$ eine Umgebung von $x$ ist und $x$ Berührpunkt von $A$, gibt es ein $a \in K(x, s)\cap A$. Dann zeigt (15.3.3), daß für $V := K(x, s)$ die Folge $(f_n|_V)_{n\in\mathbb{N}}$ gleichmäßig konvergiert. Insbesondere ist $x \in A$, dh. $U \cap \overline{A} \subseteq A$ und damit ist $A$ abgeschlossen in der Relativtopologie $\mathcal{O}_U$. Da $A$ eine nichtleere, in der Relativtopologie offene und abgeschlossene Teilmenge der zusammenhängenden Menge $U$ ist, gilt $A = U$ nach Lemma 13.5.2, dh. die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert punktweise auf ganz $U$. Mit der Abschätzung (15.3.3) folgt aber wie eben aus der punktweisen die lokal gleichmäßige Konvergenz von $(f_n)_{n\in\mathbb{N}}$, also ist $f$ nach Satz 13.4.3 stetig.

Es sei $a \in U$, $\varepsilon > 0$ und $r > 0$ so gewählt, daß $V := K(a, r) \subseteq U$ ist und $(f'_n|_V)_{n\in\mathbb{N}}$ gleichmäßig konvergiert. Dann gibt es $N = N(\varepsilon) \in \mathbb{N}$, so daß

$$\sup\{|||f'_n(z) - f'_m(z)||| : z \in V\} < \varepsilon$$

für alle $n, m \geq N(\varepsilon)$ erfüllt ist. Fixiere $n \geq N(\varepsilon)$, so gilt

$$|||g(a)-f'_n(a)||| = \left|\left|\left| \lim_{m\to\infty} f'_m(a)-f'_n(a)\right|\right|\right| \leq \sup\{|||f'_m(z)-f'_n(z)||| : z \in V, m \geq N(\varepsilon)\} \leq \varepsilon. \tag{15.3.4}$$

Betrachte nun in (15.3.3) den Limes $m \to \infty$, dann ist

$$\begin{aligned} \|f(x) - f(a) - (f_n(x) - f_n(a))\| &= \left\|\lim_{m\to\infty} (f_m(x) - f_m(a)) - (f_n(x) - f_n(a))\right\| \\ &\leq \|x - a\| \sup\{|||f'_m(z) - f'_n(z)||| : z \in V, m \geq N(\varepsilon)\} \leq \|x - a\|\varepsilon \qquad (15.3.5) \end{aligned}$$

Da $f_n$ in $a$ differenzierbar ist, gibt es ein $s \in ]0, r]$, so daß

$$\|f_n(x) - f_n(a) - f'_n(a)[x - a]\| \leq \varepsilon\|x - a\| \tag{15.3.6}$$

für alle $\|x - a\| \leq s$ gilt. Kombiniert man (15.3.4) - (15.3.6), dann folgt

$$\begin{aligned} \|f(x) - f(a) - g(a)[x - a]\| \leq\ & \|f(x) - f(a) - (f_n(x) - f_n(a))\| \\ & +\|f_n(x) - f_n(a) - f'_n(a)[x - a]\| \\ & +\|(f'_n(a) - g'(a))[x - a]\| \leq 3\varepsilon\|x - a\| \qquad (15.3.7) \end{aligned}$$

für jedes $x \in X$ mit $\|x-a\| < s$, dh. $f$ ist im Punkt $a$ differenzierbar mit $f'(a) = g(a)$. □

**Korollar 15.3.5** (gliedweises Differenzieren von Potenzreihen)**.** *Es sei $X$ ein Banachraum und $(c_n)_{n\in\mathbb{N}}$ eine Folge in $X$ und $a \in \mathbb{C}$. Es sei $\rho \in ]0, \infty]$ der Konvergenzradius der*

*Potenzreihe* $\sum_{k=0}^{\infty} c_k(z-a)^k$, *dann ist die stetige Funktion*

$$\begin{array}{rcl} f : \{z \in \mathbb{C} : |z-a| < \rho\} & \to & X \\ z & \mapsto & \sum_{k=0}^{\infty} c_k(z-a)^k \end{array}$$

*differenzierbar mit*

$$\begin{array}{rcl} f' : \{z \in \mathbb{C} : |z-a| < \rho\} & \to & X \\ z & \mapsto & \sum_{k=1}^{\infty} kc_k(z-a)^{k-1} \end{array}$$

*Beweis.* Nach Beispiel 15.1.10 sind alle Monome $z \mapsto z^k$ differenzierbar, also ist jedes Polynom differenzierbar. Insbesondere ist jede Partialsumme

$$\begin{array}{rcl} f_n : \mathbb{C} & \to & X \\ z & \mapsto & \sum_{k=0}^{n} c_k(z-a)^k \end{array}$$

differenzierbar mit Ableitung

$$\begin{array}{rcl} f_n' : \mathbb{C} & \to & X \\ z & \mapsto & \sum_{k=1}^{n} kc_k(z-a)^{k-1} \end{array}$$

Die Potenzreihe $\sum_{k=1}^{\infty} kc_k(z-a)^{k-1}$ hat denselben [4] Konvergenzradius $\rho$, wie die Potenzreihe $\sum_{k=1}^{\infty} c_k(z-a)^k$. Daher sind auf $U := \{z \in \mathbb{C} : \|z-a\| < \rho\}$ die Funktionenfolgen $(f_n|_U)_{n\in\mathbb{N}}$ und $(f_n'|_U)_{n\in\mathbb{N}}$ lokal gleichmäßig konvergent mit den Grenzwerten $f$ bzw. $f'$ und somit erfüllen diese beiden Funktionenfolgen die Voraussetzung von Satz 15.3.4, woraus die Behauptung folgt. □

**Beispiel 15.3.6.** Der natürliche Logarithmus $\ln : ]0,\infty[ \to \mathbb{R}$ ist auf $]0,\infty[$ differenzierbar mit

$$(\ln)'(x) = \frac{1}{x} \tag{15.3.8}$$

[4] $\sum_{k=1}^{\infty} kc_k(z-a)^{k-1}$ konvergiert genau dann, wenn $\sum_{k=1}^{\infty} kc_k(z-a)^k$ konvergiert und diese Potenzreihe hat mit

$$L := \overline{\lim_{k\to\infty}} \sqrt[k]{\|kc_k\|} = \overline{\lim_{k\to\infty}} \sqrt[k]{k}\, \overline{\lim_{k\to\infty}} \sqrt[k]{\|c_k\|} = \overline{\lim_{k\to\infty}} \sqrt[k]{\|c_k\|} = \frac{1}{\rho}$$

derselben Konvergenzradius.

*Beweis.* $\ln : ]0, \infty[ \to \mathbb{R}$ ist nach Definition die Umkehrfunktion von $\exp|_{\mathbb{R}} : \mathbb{R} \to ]0, \infty[$ und nach Lemma 13.8.4 stetig Nach Korollar 15.3.5 ist $\exp|_{\mathbb{R}} : \mathbb{R} \to ]0, \infty[$ differenzierbar in jedem Punkt $a \in \mathbb{R}$ mit Ableitung $\exp'(a) : \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ s & \mapsto & e^a s \end{array}$ – die wie üblich mit der Steigung $e^a$ identifiziert wird. Dann ist $\exp'(a)$ für jedes $a \in \mathbb{R}$ ein linearer Homöomorphismus, denn wegen $e^a \neq 0$ ist $\exp'(a)(\mathbb{R}) \neq \{0\}$ ein Untervektorraum von $\mathbb{R}$, also $\exp'(a)(\mathbb{R}) = \mathbb{R}$ und daher $\exp'(a)$ surjektiv, also nach Satz 8.4.4 auch bijektiv, daher nach Satz 13.9.4 ein linearer Homöomorphismus. Für $y = e^x$ lautet damit (15.1.14):

$$(\ln)'(y) = \frac{1}{(\exp)'(x)} = \frac{1}{e^x} = \frac{1}{y}.$$ □

**Lemma 15.3.7.** *Es sei $X$ ein $\mathbb{K}-$Banachraum, $S \in L(X, X)$, dann ist*

$$e^S := \sum_{n=0}^{\infty} \frac{1}{n!} S^n \in L(X, X) \tag{15.3.9}$$

*wohldefiniert. Für $S, T \in L(X, X)$ mit $TS = ST$ gilt*

$$e^{T+S} = e^T e^S \tag{15.3.10}$$

*und* $f : \begin{array}{ccc} \mathbb{R} & \to & L(X, X) \\ t & \mapsto & e^{tS} \end{array}$ *ist differenzierbar mit $f'(t) = Se^{tS}$.*

*Beweis.* Ist $X$ ein $\mathbb{K}-$Banachraum, dann ist $(L(X, X), |||\cdot|||)$ nach Satz 13.9.5 vollständig. Die Exponentialreihe hat Konvergenzradius $\infty$, daher existiert

$$e^S := \lim_{N \to \infty} \sum_{n=0}^{N} \frac{S^n}{n!} = \sum_{n=0}^{\infty} \frac{S^n}{n!}$$

in $L(X, X)$. Sind $S, T \in L(X, X)$ mit $ST = TS$, dann läßt sich auf

$$e^{S+T} = \sum_{n=0}^{\infty} \frac{(S+T)^n}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) S^k T^{n-k} = e^S e^T$$

wie beim Beweis der Funktionalgleichung der Exponentialfunktion dieser Ausdruck als Cauchyprodukt der absolut konvergenten Reihen $e^S$ und $e^T$ lesen. Gliedweises Differenzieren wie in Korollar 15.3.5 gibt $f'(t) = Se^{tS}$. □

**Bemerkung 15.3.8.** Ist $\mathcal{H}$ ein Hilbertraum und $H \in L(\mathcal{H}, \mathcal{H})$ eine stetige lineare Abbildung, dann ist zu $\psi \in \mathcal{H}$ die Funktion $\Psi : \begin{array}{ccc} \mathbb{R} & \to & \mathcal{H} \\ t & \mapsto & e^{-iHt}\psi \end{array}$ eine Lösung der Schrödingergleichung $\Psi' = -iH\Psi$. Die Voraussetzung $H \in L(\mathcal{H}, \mathcal{H})$ ist für endlich dimensionale $\mathcal{H}$ erfüllt und auch für einige Modelle auf einem Gitter also zB. für den diskreten Laplace auf $l^2(\mathbb{Z}^d)$, aber zB. nicht für den Hamiltonoperator $H$ zum Wasserstoffatom. Auch in diesem Fall ist $\Psi : \begin{array}{ccc} \mathbb{R} & \to & \mathcal{H} \\ t & \mapsto & e^{-iHt}\psi \end{array}$ Lösung der Schrödingergleichung, jedoch braucht man andere Methoden um $e^{\cdot -iHt}$ zu definieren.

**Bemerkung 15.3.9.**

a) Da jede lineare Abbildung zwischen endlichdimensionalen Vektorräumen stetig ist, gilt

$$L(\mathbb{K}^d, \mathbb{K}^d) \sim M_d(\mathbb{K})$$

indem man $S \in L(\mathbb{K}^d, \mathbb{K}^d)$ mit der darstellenden Matrix bzgl. einer Basis identifiziert.

b) Da auf dem endlichdimensionalen Vektorraum $L(\mathbb{K}^d, \mathbb{K}^d)$ alle Normen äquivalent sind, konvergiert $\sum_{n=0}^{\infty} \frac{S^n}{n!}$ für $S \in L(\mathbb{K}^d, \mathbb{K}^d)$ nicht nur in der Operatornorm $|||\cdot|||$, sondern in jeder Norm auf $L(\mathbb{K}^d, \mathbb{K}^d)$.

c) Zur konkreten Berechnung von $e^S$ ist es oft vorteilhaft die Jordanform der darstellenden Matrix zu verwenden.

## 15.4 Mittelwertsatz

**Definition 15.4.1.** *Es sei $X$ ein topologischer Raum und $f : X \to \mathbb{R}$, dann heißt $a \in X$ ein* **lokales Maximum** *bzw.* **lokales Minimum** *von $f$, wenn es eine Umgebung $V$ von $a$ gibt, so daß*

$$f(x) \leq f(a) \qquad \text{bzw.} \qquad f(x) \geq f(a) \tag{15.4.1}$$

*für alle $x \in V$ gilt. $a$ heißt* **lokales Extremum** *von $f$, wenn $a$ lokales Maximum oder lokales Minimum von $f$ ist.*

**Lemma 15.4.2.** *Es sei $U \subseteq \mathbb{R}$ offen, $f : U \to \mathbb{R}$ stetig und differenzierbar in $a \in U$ und $a$ sei ein lokales Extremum von $f$. Dann ist $f'(a) = 0$.*

*Beweis.* Wir betrachten den Fall eines lokalen Maximums: Es sei $V \subseteq U$ eine Umgebung von $a$, so daß $f(x) \leq f(a)$ für alle $x \in V$. Dann gilt $\frac{f(x) - f(a)}{x - a} \leq 0$ für alle $x \in V, x > a$ und $\frac{f(x) - f(a)}{x - a} \geq 0$ für alle $x \in V, x < a$. Da $f$ in $a$ differenzierbar ist, existiert $f'(a) = \lim\limits_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}$ nach Satz 15.1.9. Daher existieren insbesondere

$$f'_-(a) := \lim_{\substack{x \to a \\ x < a}} \frac{f(x) - f(a)}{x - a} \qquad \text{und} \qquad f'_+(a) := \lim_{\substack{x \to a \\ x > a}} \frac{f(x) - f(a)}{x - a}$$

und es gilt $f'_-(a) = f'_+(a) = f'(a)$; wegen $f'_-(a) \geq 0$ und $f'_+(a) \leq 0$ folgt $f'(a) = 0$. □

**Bemerkung 15.4.3.** Auch wenn $f : U \to \mathbb{R}$ in $a \in U$ differenzierbar mit $f'(a) = 0$ ist, so braucht $a$ kein lokales Extremum von $f$ sein, zB. für $\begin{array}{rcl} f : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^3 \end{array}$ ist $f'(0) = 0$ aber für $x < 0 < y$ ist $f(x) = x^3 < 0 = f(0) < y^3 = f(y)$, also 0 kein lokales Extremum.

**Satz 15.4.4** (Mittelwertsatz). *Es sei $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$ und $I = [\alpha, \beta]$ ein kompaktes Intervall, $f : [\alpha, \beta] \to \mathbb{R}$ stetig und auf $]\alpha, \beta[$ differenzierbar. Dann gibt es (mindestens) ein $\xi \in ]\alpha, \beta[$ mit*

$$f'(\xi) = \frac{f(\beta) - f(\alpha)}{\beta - \alpha} \tag{15.4.2}$$

*Beweis.* Die Funktion $\begin{array}{rcl} h : [\alpha, \beta] & \to & \mathbb{R} \\ x & \mapsto & f(x) - \frac{f(\beta) - f(\alpha)}{\beta - \alpha}(x - \alpha) - f(\alpha) \end{array}$ ist stetig auf $[\alpha, \beta]$ und differenzierbar auf $]\alpha, \beta[$ mit $h'(x) = f'(x) - \frac{f(\beta) - f(\alpha)}{\beta - \alpha}$ und $h(\alpha) = h(\beta) = 0$. $h$ nimmt nach Lemma 13.6.8 auf der kompakten Menge $[\alpha, \beta]$ Maximum und Minimum an

- Ist $\alpha$ Maximum und $\beta$ Minimum von $h$ oder ist $\alpha$ Minimum und $\beta$ Maximum von $h$, so ist $h$ konstant auf $[\alpha, \beta]$ wegen $h(\alpha) = h(\beta) = 0$. Daher ist $h'(\xi) = 0$ für alle $\xi \in ]\alpha, \beta[$ und wegen $h'(\xi) = f'(\xi) - \frac{f(\beta) - f(\alpha)}{\beta - \alpha}$ folgt (15.4.2) in diesem Fall.
- Andernfalls hat $h$ in einem $\xi \in ]\alpha, \beta[$ ein Extremum und da $h$ auf $]\alpha, \beta[$ differenzierbar ist, folgt $h'(\xi) = 0$, dh. $h'(\xi) = f'(\xi) - \frac{f(\beta) - f(\alpha)}{\beta - \alpha} = 0$ und (15.4.2) folgt auch in diesem Fall. □

**Bemerkung 15.4.5.** Der Mittelwertsatz gilt in dieser Form nur für reellwertige Funktionen, aber schon nicht mehr für komplexwertige Funktionen. Beispiel: $\begin{array}{rcl} f : \mathbb{R} & \to & \mathbb{C} \\ t & \mapsto & e^{it} \end{array}$ ist $2\pi$-periodisch, also etwa $f(2\pi) - f(0) = 0$ und nach Korollar 15.3.5 ist $f'(t) = ie^{it}$, also $f'(t) \neq 0$ für alle $t \in \mathbb{R}$ nach Satz 6.6.3.

**Korollar 15.4.6.** *Ist $f : ]\alpha, \beta[ \to \mathbb{R}$ differenzierbar, so gilt:*

- *Ist $f'(x) > 0$ für alle $x \in ]\alpha, \beta[$, so ist $f$ streng monoton steigend in $]\alpha, \beta[$.*
- *$f$ ist genau dann monoton steigend in $]\alpha, \beta[$, wenn $f'(x) \geq 0$ für alle $x \in ]\alpha, \beta[$ ist.*

*Beweis.* Für $\alpha < x_1 < x_2 < \beta$ gibt es nach dem Mittelwertsatz ein $\xi \in ]x_1, x_2[$ mit $f(x_2) - f(x_1) = (x_2 - x_1)f'(\xi)$, dh. aus $f'(\xi) > 0$ folgt $f(x_2) > f(x_1)$. Ebenso folgt $f(x_2) \geq f(x_1)$ falls $f'(x) \geq 0$ für alle $x \in ]\alpha, \beta[$ ist. Ist $f : ]\alpha, \beta[ \to \mathbb{R}$ monoton steigend, dann ist $\frac{f(a+h) - f(a)}{h} \geq 0$ für jedes $a \in ]\alpha, \beta[$ und $h \in \mathbb{R} \setminus \{0\}$ mit [5] $a + h \in ]\alpha, \beta[$. Da $f$ in $a \in ]\alpha, \beta[$ differenzierbar ist, existiert $f'(a) = \lim\limits_{\substack{h \to 0 \\ h \neq 0}} \frac{f(a+h) - f(a)}{h}$ und daher ist $f'(a) \geq 0$. □

[5] wie man durch Fallunterscheidung

- $h > 0$, dann ist $f(a+h) - f(a) \geq 0$
- $h < 0$, dann ist $f(a+h) - f(a) \leq 0$

zeigt

**Lemma 15.4.7** (verallgemeinerter Mittelwertsatz)**.** *Es seien $a, b \in \widehat{\mathbb{R}}$ mit $a < b$. $f : [a, b] \to \mathbb{R}$ und $g : [a, b] \to \mathbb{R}$ seien stetig und auf $]a, b[$ differenzierbar und es sei $g'(x) \neq 0$ für alle $x \in ]a, b[$ und es sei $g(a) \neq g(b)$ im Fall $a \notin \mathbb{R}$ oder $b \notin \mathbb{R}$. Dann [6] gibt es ein $\xi \in ]a, b[$ mit*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}. \tag{15.4.3}$$

*Beweis.* Nach dem Mittelwertsatz folgt im Fall $a, b \in \mathbb{R}$ dann $g(a) \neq g(b)$ aus $g'(x) \neq 0$ für alle $x \in ]a, b[$, bzw. dies gilt laut Voraussetzung im Fall $a \notin \mathbb{R}$ oder $b \notin \mathbb{R}$. Somit ist

$$\begin{array}{rcl} h : [a, b] & \to & \mathbb{R} \\ x & \mapsto & f(x) - \dfrac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)) \end{array}$$

eine wohldefinierte, auf $[a, b]$ stetige und auf $]a, b[$ differenzierbare Funktion mit $h(a) = h(b)$. Da $[a, b]$ kompakt ist (falls $a \notin \mathbb{R}$ oder $b \notin \mathbb{R}$ als Teilmenge von $\widehat{\mathbb{R}}$) besitzt $h$ ein Maximum und ein Minimum und das analoge Argument wie im Beweis des Mittelwertsatzes zeigt, daß es ein $\xi \in ]a, b[$ mit

$$0 = h'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(\xi)$$

gibt, woraus wegen $g'(\xi) \neq 0$ die Behauptung folgt. □

**Lemma 15.4.8** (l' Hospital)**.** *Es sei $a, b \in \widehat{\mathbb{R}}$ mit $a < b$ und $f : ]a, b[ \to \mathbb{R}$ und $g : ]a, b[ \to \mathbb{R}$ seien differenzierbar mit $g'(x) \neq 0$ für alle $x \in ]a, b[$. Ist $\lim\limits_{x \searrow a} f(x) = \lim\limits_{x \searrow a} g(x) = 0$ oder $\lim\limits_{x \searrow a} f(x) = \lim\limits_{x \searrow a} g(x) = \infty$, so gilt: Existiert $\lim\limits_{x \searrow a} \dfrac{f'(x)}{g'(x)}$ in $\mathbb{R}$, dann existiert auch $\lim\limits_{x \searrow a} \dfrac{f(x)}{g(x)}$ und es gilt:*

$$\lim_{x \searrow a} \frac{f'(x)}{g'(x)} = \lim_{x \searrow a} \frac{f(x)}{g(x)}.$$

*Analoge Aussagen gelten auch für den Fall $x \nearrow b$.*

*Beweis.* Wir betrachten den Fall $\lim\limits_{x \searrow a} f(x) = \lim\limits_{x \searrow a} g(x) = 0$, dann können wir die Funktionen $f$ und $g$ stetig fortsetzen zu

$$\begin{array}{rcl} F : [a, b[ & \to & \mathbb{R} \\ x & \mapsto & F(x) := \begin{cases} f(x) & \text{für } x \in ]a, b[ \\ 0 & \text{für } x = a \end{cases} \end{array}$$

und

$$\begin{array}{rcl} G : [a, b[ & \to & \mathbb{R} \\ x & \mapsto & G(x) := \begin{cases} g(x) & \text{für } x \in ]a, b[ \\ 0 & \text{für } x = a \end{cases} \end{array}$$

. Nach dem verallgemeinerten Mittelwertsatz gibt es dann für jedes $x \in ]a, b[$ ein $\xi \in ]a, x[$ mit

$$\frac{f(x)}{g(x)} = \frac{F(x) - F(a)}{G(x) - G(a)} = \frac{F'(\xi)}{G'(\xi)} = \frac{f'(\xi)}{g'(\xi)} \tag{15.4.4}$$

[6] Beachte, daß hier auch der Fall $a = -\infty$ und $b = \infty$ zugelassen ist; da $f$ und $g$ reellwertig und $g(a) \neq g(b)$ sind, ist der Differenzenquotient in (15.4.3) auch dann definiert

Hier impliziert $x \searrow a$ auch $\xi \searrow a$, daher folgt aus der Existenz von $\lim\limits_{\xi \searrow a} \frac{f'(\xi)}{g'(\xi)} \in \mathbb{R}$ mit (15.4.4) die Gleichheit $\lim\limits_{x \searrow a} \frac{f'(x)}{g'(x)} = \lim\limits_{x \searrow a} \frac{f(x)}{g(x)}$. $\square$

## 15.5 Richtungsableitungen, partielle Ableitungen

**Bemerkung 15.5.1.** Es seien $X$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen mit $a \in U$, $f : U \to Y$ und $v \in X \backslash \{\mathbf{0}\}$. Da $U$ offen ist, gibt es ein $\varepsilon > 0$ mit $K(a, \varepsilon) := \{x \in X : \|x - a\| < \varepsilon\} \subseteq U$, daher gilt $a + tv \in K(a, \varepsilon) \subseteq U$ für $r := \frac{\varepsilon}{\|v\|}$ und jedes $t \in \mathbb{K}$ mit $|t| < r$. Deshalb ist die Funktion

$$\begin{array}{rcl} \widetilde{f}_v : \{t \in \mathbb{K} : |t| < r\} & \to & Y \\ t & \mapsto & f(a + tv) \end{array}$$

wohldefiniert.

**Definition 15.5.2.** *Ist in der Situation von Bemerkung 15.5.1 die Funktion $\widetilde{f}_v$ an der Stelle $t = 0$ differenzierbar, so heißt die Ableitung $(D\widetilde{f}_v)(0)$ von $\widetilde{f}_v$ bei $0$ die* **Richtungsableitung von $f$ an der Stelle $a$ in Richtung $v$.** *Schreibweise: $(D_v f)(a)$.*

**Lemma 15.5.3.** *Es seien $X$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen und $f : U \to Y$ differenzierbar in $a \in U$. Dann existiert für jedes $v \in X \backslash \{\mathbf{0}\}$ die Richtungsableitung $(D_v f)(a)$ an der Stelle $a$ in Richtung $v$ und es gilt:*

$$(D_v f)(a) = (Df)(a)[v]. \tag{15.5.1}$$

*Beweis.* Für $v \in X \backslash \{\mathbf{0}\}$ und $a \in U$ wähle $\varepsilon, r > 0$ wie in Bemerkung 15.5.1, dann ist $a + tv \in U$ für alle $t \in \mathbb{K}$ mit $|t| < r$ und wegen der Differenzierbarkeit von $f$ in $a$ gibt es nach Lemma 15.1.7 eine in $a$ stetige Funktion $r_a : U \to Y$ mit $r_a(a) = \mathbf{0}$ und

$$f(a + tv) = f(a) + (Df)(a)[tv] + r_a(a + tv)\|tv\| = f(a) + t(Df)(a)[v] + r_a(a + tv)\|tv\|$$

für alle $a + tv \in U$. Wegen $(Df)(a)[v] \in Y$ definiert $\begin{array}{rcl} \{t \in \mathbb{K} : |t| < r\} & \to & Y \\ t & \mapsto & t(Df)(a)[v] \end{array}$ ein Element in $L(\mathbb{K}, Y)$ und für $\begin{array}{rcl} g : \{t \in \mathbb{K} : |t| < r\} & \to & Y \\ t & \mapsto & f(a + tv) \end{array}$ ist

$$\lim_{\substack{t \to 0 \\ t \neq 0}} \frac{g(t) - g(0)}{t} = \lim_{\substack{t \to 0 \\ t \neq 0}} \left( \frac{t(Df)(a)[v]}{t} + \frac{|t| r_a(tv) \|v\|}{t} \right) = (Df)(a)[v]$$

$\square$

**Bemerkung 15.5.4.** Die Umkehrung der Aussage von Lemma 15.5.3 gilt nicht, wie das folgende Beispiel mit nur zwei reellen Variablen zeigt:

$$\begin{array}{rcl} f : \mathbb{R}^2 & \to & \mathbb{R} \\ (x, y) & \mapsto & \begin{cases} \dfrac{x^2 y}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases} \end{array}$$

Denn für jedes $v = (v_1, v_2) \in \mathbb{R}^2 \backslash \{(0,0)\}$ und $t \neq 0$ ist $f(tv) = \dfrac{t^3 v_1^2 v_2}{t^2(v_1^2 + v_2^2)} = tf(v)$, woraus

$$(D_v f)(\mathbf{0}) = \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(tv) - f(\mathbf{0})}{t} = f(v)$$

folgt, was zeigt, daß in $(0,0)$ jede Richtungsableitung $(D_v f)(\mathbf{0})$ existiert. Aber $f$ ist in $\mathbf{0}$ nicht differenzierbar, denn sonst folgt $(Df)(\mathbf{0})[v] = (D_v f)(\mathbf{0}) = f(v)$ für alle $v \in \mathbb{R}^2$ aus Lemma 15.5.3. Das kann aber nicht sein, denn $\begin{array}{rcl} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ v & \mapsto & (Df)(\mathbf{0})[v] \end{array}$ ist linear, aber $f$ nicht.

**Lemma 15.5.5.** *Es sei $X$ ein Banachraum, $U \subseteq X$ offen, $a \in U$ sei ein lokales Extremum von $f : U \to \mathbb{R}$ und in $a$ existiere für jedes $v \in X \backslash \{\mathbf{0}\}$ die Richtungsableitung $(D_v f)(a)$. Dann gilt*

$$(D_v f)(a) = 0 \quad \text{für alle } v \in X \backslash \{\mathbf{0}\}. \tag{15.5.2}$$

*Beweis.* Zu $v \in X \backslash \{\mathbf{0}\}$ wähle $r > 0$ und $\begin{array}{rcl} \widetilde{f_v} : \{t \in \mathbb{K} : |t| < r\} & \to & \mathbb{R} \\ t & \mapsto & f(a + tv) \end{array}$ wie in Bemerkung 15.5.1. Dann ist $\widetilde{f_v}$ in 0 differenzierbar und da $f$ in $a$ ein lokales Extremum besitzt, hat $\widetilde{f_v}$ ein Extremum bei $t = 0$ und daher ist $\widetilde{f_v}'(a) = 0$ nach Lemma 15.4.2. □

**Bemerkung 15.5.6.** Es seien $X = X_1 \times ... \times X_n$ und $Y$ $\mathbb{K}$-Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen und $a = (a_1, ..., a_n) \in U$. Dann ist für jedes $j \in \{1, ..., n\}$ die Menge

$$U_j(a) := \{x_j \in X_j : (a_1, ..., a_{j-1}, x_j, a_{j+1}, ..., a_n) \in U\}$$

eine offene Umgebung von $a_j$ in $X_j$.

*Beweis.* Zu $x_j \in U_j(a)$ ist $(a_1, ..., a_{j-1}, x_j, a_{j+1}, ..., a_n) \in U$, daher gibt es ein $r > 0$, so daß für jedes $y = (y_1, ..., y_n) \in X_1 \times ... \times X_n$ mit $\|(a_1, ..., a_{j-1}, x_j, a_{j+1}, ..., a_n) - y\|_X < r$ gilt: $y \in U$. Dies gilt insbesondere für jede Wahl von $y = (a_1, ..., a_{j-1}, y_j, a_{j+1}, ..., a_n) \in X$ mit $\|x_j - y_j\| < r$, woraus $\{y_j \in X_j : \|y_j - x_j\| < r\} \subseteq U_j(a)$ folgt, was zeigt, daß $U_j(a)$ offen in $X_j$ ist. □

**Definition 15.5.7.** *Es seien $X = X_1 \times ... \times X_n$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen und $a = (a_1, ..., a_n) \in U$. Eine stetige Funktion $f : U \to Y$ heißt bezüglich der $j$-ten Variablen* **partiell differenzierbar** *in $a$, wenn*

$$\begin{array}{rcl} h_j : U_j(a) & \to & Y \\ x_j & \mapsto & f(a_1, ..., a_{j-1}, x_j, a_{j+1}, ..., a_n) \end{array}$$

*im Punkt $a_j$ differenzierbar ist. $h_j'(a_j) \in L(X_j, Y)$ heißt die* **partielle Ableitung** *von $f$ in $a$ bezüglich der $j$-ten Variablen. Schreibweise: $(D_j f)(a)$.*

**Bemerkung 15.5.8.** Ist $\dim X_1 = ... = \dim X_n = 1$, also oE. $X_1 = ... = X_n = \mathbb{K}$ und $U \subseteq \mathbb{K}^n$ offen, $a = (a_1, ..., a_n) \in U$, so stimmt die partielle Ableitung $(D_j f)(a)$ von $f : U \to Y$ bezüglich der $j$-ten Variablen in $a$ mit der Richtungsableitung $(D_{\underline{e}_j} f)(a)$ in Richtung des $j$-ten Einheitsvektors $\underline{e}_j$ überein. In diesem Fall schreiben wir $\frac{\partial f}{\partial x_j}(a)$ statt $(D_{\underline{e}_j} f)(a)$

**Satz 15.5.9.** *Es seien $X = X_1 \times ... \times X_n$ und $Y$ $\mathbb{K}$-Banachräume, $U \subseteq X$ offen und $f : U \to Y$ stetig. Dann sind äquivalent:*

a) *$f$ ist **stetig differenzierbar** auf $U$, dh.* $\begin{array}{rcl} Df : U & \to & L(X,Y) \\ a & \mapsto & (Df)(a) \end{array}$ *existiert und ist stetig.*

b) *In jedem Punkt $a \in U$ ist für alle $j \in \{1, ..., n\}$ die Funktion $f$ partiell differenzierbar bezüglich der $j$-ten Variablen und die partiellen Ableitungen*

$$\begin{array}{rcl} D_1 f : U & \to & L(X_1, Y) \\ x & \mapsto & (D_1 f)(x) \end{array}, ..., \begin{array}{rcl} D_n f : U & \to & L(X_n, Y) \\ x & \mapsto & (D_n f)(x) \end{array}$$

*sind stetig – dazu sagt man kurz **stetig partiell differenzierbar auf** $U$.*

*In diesem Fall ist die Ableitung $(Df)(a) \in L(X,Y)$ von $f$ im Punkt $a$ gegeben als*

$$(Df)(a)[x_1, ..., x_n] = \sum_{j=1}^{n} (D_j f)(a)[x_j] \tag{15.5.3}$$

*Beweis.*

a) $\Rightarrow$ b) Es sei $f$ stetig differenzierbar auf $U$, dann ist für jedes $j \in \{1, ..., n\}$ und für $a = (a_1, ..., a_n) \in U$ die Abbildung

$$\begin{array}{rclcl} h_j : U_j(a) & \xrightarrow{i_j} & U & \xrightarrow{f} & Y \\ x_j & \longmapsto & (a_1, ..., a_{j-1}, x_j, a_{j+1}, ..., a_n) & \longmapsto & f(a_1, ..., a_{j-1}, x_j, a_{j+1}, ..., a_n) \end{array}$$

die Hintereinanderausführung von

$$\begin{array}{rcl} i_j : U_j(a) & \to & U \\ x_j & \mapsto & (a_1, ..., a_{j-1}, x_j, a_{j+1}, ..., a_n) \end{array}$$

und $f$. Nach Satz 15.1.12 ist $i_j$ differenzierbar, denn die konstanten Abbildungen $x_j \mapsto a_k$ mit $k \neq j$ sind differenzierbar mit Ableitung 0 und auch $x_j \mapsto x_j$ ist als Einschränkung der identischen Abbildung differenzierbar, also ist

$$\begin{array}{rcl} i_j'(b) : X_j & \to & X \\ x_j & \mapsto & (0, ..., 0, x_j, 0, ..., 0) \end{array} \in L(X_j, X)$$

für jedes $b \in U_j(a)$, damit ist $\begin{array}{rcl} i_j' : U_j(a) & \to & L(X_j, X) \\ b & \mapsto & i_j'(b) \end{array}$ konstant, also insbesondere stetig. Nach Satz 15.1.11 ist

$$h_j'(x_j) = (Df)(a_1, ..., a_{j-1}, x_j, a_{j+1}, ..., a_n) \circ (Di_j)(x_j),$$

dh. $h_j$ ist differenzierbar auf $U_j(a)$ und $h_j'$ ist stetig als Verknüpfung stetiger Funktionen und somit ist $f$ bezüglich jeder Variablen stetig partiell differenzierbar auf $U$.

b) $\Rightarrow$ a) Es sei $f$ stetig partiell differenzierbar auf $U$ und $a = (a_1, ..., a_n) \in U$; da $U$ offen ist, gibt es ein $r > 0$, so daß für alle $x = (x_1, ..., x_n) \in X$ mit $\|x\| < r$ dann $a + x \in U$ ist.

$$\begin{aligned} & f(a_1 + x_1, ..., a_n + x_n) - f(a_1, ..., a_n) - \sum_{j=1}^{n} (D_j f)(a_1, ..., a_n)[x_j] \\ = & \sum_{j=1}^{n} \Big( f(a_1 + x_1, ..., a_j + x_j, a_{j+1}, ..., a_n) - f(a_1 + x_1, ..., a_{j-1} + x_{j-1}, a_j, ..., a_n) \\ & \qquad - (D_j f)(a_1 + x_1, ..., a_{j-1} + x_{j-1}, a_j, ..., a_n)[x_j] \Big) \\ & + \sum_{j=2}^{n} \Big( (D_j f)(a_1 + x_1, ..., a_{j-1} + x_{j-1}, a_j, ..., a_n)[x_j] - (D_j f)(a_1, ..., a_n)[x_j] \Big) \end{aligned}$$

Da jede partielle Ableitung $D_j f$, $j = 1, ..., n$ stetig ist, gibt es für jedes $\varepsilon > 0$ ein $r(\varepsilon) > 0$ mit

$$|||(D_j f)(a_1 + x_1, ..., a_{j-1} + x_{j-1}, a_j, ..., a_n) - (D_j f)(a_1, ..., a_n)||| \leq \varepsilon$$

für $\|x_1\|, ..., \|x_{j-1}\| < r(\varepsilon)$ und nach Definition der partiellen Ableitung gilt – nach evtl. Verkleinern von $r(\varepsilon)$ für $\|x_1\|, ..., \|x_{j-1}\| < r(\varepsilon)$:

$$\begin{aligned} & \Big\| f(a_1 + x_1, ..., a_j + x_j, a_{j+1}, ..., a_n) - f(a_1 + x_1, ..., a_{j-1} + x_{j-1}, a_j, ..., a_n) - \\ & \quad (D_j f)(a_1 + x_1, ..., a_{j-1} + x_{j-1}, a_j, ..., a_n)[x_j] \Big\| \leq \varepsilon \|x_j\|. \end{aligned}$$

Für $\|x_1\|, ..., \|x_n\| \leq r(\varepsilon)$ ist dann

$$\frac{\left\| f(a_1 + x_1, ..., a_n + x_n) - f(a_1, ..., a_n) - \sum_{j=1}^{n} (D_j f)(a_1, ..., a_n)[x_j] \right\|}{\|(x_1, ..., x_n)\|} \leq \varepsilon,$$

Daher ist $f$ in $a = (a_1, .., a_n)$ differenzierbar und $f'(a) \in L(X, Y)$ ist gegeben als

$$\begin{aligned} f'(a_1, ..., a_n) : X & \to Y \\ (x_1, ..., x_n) & \mapsto \sum_{j=1}^{n} (D_j f)(a_1, ..., a_n)[x_j] \end{aligned}$$

und da die partiellen Ableitungen stetig sind, ist auch $f'$ stetig. □

**Bemerkung 15.5.10.** Wir beim Beweis von Satz 15.5.9 bei der Folgerung a) $\Rightarrow$ b) zeigt man, daß jede in $a$ differenzierbare Funktion $f$ in $a$ bezüglich jeder der Variablen partiell differenzierbar ist und

$$(Df)(a)[x_1, ..., x_n] = \sum_{j=1}^{n} (D_j f)(a)[x_j] \tag{15.5.4}$$

gilt. Ohne die Stetigkeit der partiellen Ableitung ist die Umkehrung falsch, dazu reicht das Beispiel

$$\begin{array}{rcl} f:\mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & \begin{cases} 0 & \text{für } (x,y)=(0,0) \\ \dfrac{x^2y}{x^2+y^2} & \text{für } (x,y)\neq(0,0) \end{cases} \end{array}$$

aus Bemerkung 15.5.4.

**Definition 15.5.11.** *Es sei $n\in\mathbb{N}$, $U\subseteq\mathbb{R}^n$ offen, $a\in U$ und $f:U\to\mathbb{R}$ sei differenzoerbar in $a\in U$. Ist $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf $\mathbb{R}^n$, so gibt es nach dem Rieszschen Darstellungssatz* [7] *ein eindeutig bestimmtes Element* $(\mathrm{grad} f)(a)\in\mathbb{R}^n$, *so daß sich die lineare Abbildung* $\begin{array}{rcl} f'(a):\mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & f'(a)[x] \end{array}$ *in der Form* $\begin{array}{rcl} f'(a):\mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & f'(a)[x]=\langle(\mathrm{grad} f)(a),x\rangle \end{array}$ *schreibt.* $(\mathrm{grad} f)(a)$ *heißt der* **Gradient** *von $f$ an der Stelle $a$.*

**Lemma 15.5.12.** *Es sei $U\subseteq\mathbb{R}^n$ offen und $f:U\to\mathbb{R}$ stetig partiell differenzierbar. Ist $(\underline{e}_1,\dots,\underline{e}_n)$ die Standardbasis von $\mathbb{R}^n$ und $\langle\cdot,\cdot\rangle$ das Standardskalarprodukt auf $\mathbb{R}^n$, dann ist (in der Standardbasis):*

$$(\mathrm{grad} f)(a)=\begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a)\end{pmatrix}=\sum_{k=1}^n \underline{e}_k\frac{\partial f}{\partial x_k}(a). \tag{15.5.5}$$

*Beweis.* Ist $f$ stetig partiell differenzierbar, so ist $f$ nach Satz 15.5.9 stetig differenzierbar und

$$(Df)(a)[x_1,\dots,x_n]=\sum_{k=1}^n\frac{\partial f}{\partial x_k}(a)[x_k]=\left\langle\begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a)\end{pmatrix},\begin{pmatrix} x_1\\ \vdots \\ x_n\end{pmatrix}\right\rangle$$

gilt für alle $\underline{x}=\begin{pmatrix} x_1\\ \vdots \\ x_n\end{pmatrix}\in\mathbb{R}^n$. Aus Lemma 5.4.15 folgt dann $(\mathrm{grad} f)(a)=\begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a)\end{pmatrix}$. □

**Bemerkung 15.5.13.** Es sei $U\subseteq\mathbb{R}^n$ offen, $f:U\to\mathbb{R}$ stetig partiell differenzierbar und $a\in U$ mit $(\mathrm{grad} f)(a)\neq\mathbf{0}$. Dann gibt $(\mathrm{grad} f)(a)$ die Richtung des **steilsten Anstiegs** von $f$ in $a$ an, dh. $(\mathrm{grad} f)(a)$ ist die Richtung $v$, in der die Richtungsableitung $(D_vf)(a)$ den maximalen Wert annimmt.

---

[7] siehe Kapitel "Hilberträume". Wir haben bereits bei der Definition des Skalarprodukts gesehen, daß für jedes $b\in\mathbb{R}^n$ die Abbildung $\begin{array}{rcl}\langle b,\cdot\rangle:\mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & \langle b,x\rangle\end{array}$ eine lineare Abbildung ist. Der Rieszsche Darstellungsatz sagt, daß sich jede lineare Abbildung $l:\mathbb{R}^n\to\mathbb{R}$ in dieser Form schreiben läßt.

*Beweis.* Setze $w := \dfrac{(\operatorname{grad} f)(a)}{\|(\operatorname{grad} f)(a)\|}$, so ist $\|w\| = 1$ und damit nach der Wahl von $w$:

$$(Df)(a)[w] = \langle (\operatorname{grad} f)(a), w\rangle = \frac{\langle (\operatorname{grad} f)(a), (\operatorname{grad} f)(a)\rangle}{\|(\operatorname{grad} f)(a)\|} = \|(\operatorname{grad} f)(a)\|. \tag{15.5.6}$$

Nach Definition ist die Operatornorm

$$|||\langle (\operatorname{grad} f)(a), \cdot\rangle||| = \sup\{|\langle (\operatorname{grad} f)(a), z\rangle| : \|z\| = 1\} = \|(\operatorname{grad} f)(a)\| \tag{15.5.7}$$

Da $\{z \in \mathbb{R}^n : \|z\| = 1\}$ kompakt und $\langle (\operatorname{grad} f)(a), \cdot\rangle : \mathbb{R}^n \to \mathbb{R}$ stetig ist, ist das Supremum ein Maximum, also

$$\sup\{|\langle (\operatorname{grad} f)(a), z\rangle| : z \in \mathbb{R}^n, \|z\| = 1\} = \max\{|\langle (\operatorname{grad} f)(a), z\rangle| : z \in \mathbb{R}^n, \|z\| = 1\}$$

und damit ist nach Lemma 15.5.3

$$\begin{aligned}(D_w f)(a) &= (Df)(a)[w] = \|(\operatorname{grad} f)(a)\| = \max\{|\langle (\operatorname{grad} f)(a), z\rangle| : z \in \mathbb{R}^n, \|z\| = 1\} \\ &= \max\{|(D_z f)(a)| : z \in \mathbb{R}^n, \|z\| = 1\}\end{aligned}$$ □

**Definition 15.5.14.** *Es sei $U \subseteq \mathbb{K}^n$ offen und $a \in U$ und*

$$\begin{aligned} f = (f_1, ..., f_m) : U &\to \mathbb{K}^m \\ x &\mapsto \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}\end{aligned}$$

*partiell differenzierbar. Dann heißt*

$$(Jf)(a) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} \tag{15.5.8}$$

*die* **Jacobimatrix** *von $f$ in $a$.*

**Lemma 15.5.15.** *Es sei $U \subseteq \mathbb{K}^n$ offen, $a \in U$ und $f = (f_1, ..., f_m) : U \to \mathbb{K}^m$ in $a$ differenzierbar, dann ist jede der Koeffizientenfunktionen $f_1, ..., f_m$ partiell differenzierbar in $a$ und $(Jf)(a)$ ist die darstellende Matrix der Ableitung $(Df)(a)$ von $f$ in $a$ bezüglich der Standardbasen.*

*Beweis.* Nach Satz 15.1.12 ist $f$ genau dann in $a$ differenzierbar, wenn jede der Koordinatenfunktionen $f_j = \operatorname{pr}_j \circ f$, $j = 1, ..., m$ in $a$ differenzierbar ist. Nach Bemerkung 15.5.10 folgt die partielle Differenzierbarkeit von $f_j$ in $a$. Ist $(\alpha_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in M(m \times n, \mathbb{K})$ die darstellende Matrix der linearen Abbildung $(Df)(a)$ bezüglich der Standardbasen, so bedeutet dies

$$(Df)(a)[\underline{e}_k] = \sum_{l=1}^{m} \alpha_{lk} \underline{e}_l$$

für $k = 1, ..., n$. Nach Satz 15.1.12 und Lemma 15.5.3 ist

$$(Df)(a)[\underline{e}_k] = \begin{pmatrix} (Df_1)(a)[\underline{e}_k] \\ \vdots \\ (Df_m)(a)[\underline{e}_k] \end{pmatrix} = \begin{pmatrix} (D_{\underline{e}_k} f_1)(a) \\ \vdots \\ (D_{\underline{e}_k} f_m)(a) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_k}(a) \\ \vdots \\ \frac{\partial f_m}{\partial x_k}(a) \end{pmatrix} = \sum_{l=1}^{m} \frac{\partial f_l}{\partial x_k}(a)\underline{e}_l$$

dh. $\alpha_{lk} = \frac{\partial f_l}{\partial x_k}(a)$, da $\underline{e}_1, ..., \underline{e}_m$ eine Basis von $\mathbb{K}^m$ ist. □

**Bemerkung 15.5.16.** Da jede lineare Abbildung $l : \mathbb{K} \to \mathbb{K}$ die Form $z \mapsto az$ hat, bekommen wir für jedes $A = (a) = (\alpha + i\beta) \in M_1(\mathbb{C})$ die Formel

$$A[z] = az = (\alpha + i\beta)(x + iy) = \alpha x - \beta y + i(\alpha y + \beta x).$$

Nach der Identifikation von $\mathbb{C}$ mit $\mathbb{R}^2$ durch $\begin{array}{rcl} i : \mathbb{C} & \to & \mathbb{R}^2 \\ z = x + iy & \mapsto & \begin{pmatrix} x \\ y \end{pmatrix} \end{array}$ bekommen wir

$$i(A[z]) = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Daher ist nur jedes $A \in M_2(\mathbb{R})$ der Form $A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$ (nach Identifikation mittels $i$) die darstellende Matrix einer $\mathbb{C}$-linearen Abbildung. Ist also $\begin{array}{rcl} f : \mathbb{R}^2 & \to & \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & \begin{pmatrix} u(x,y) \\ v(x,y) \end{pmatrix} \end{array}$ stetig reell partiell differenzierbar, dann definiert (nach Identifikation mit $i$) die Jacobimatrix $(Jf)(a) = \begin{pmatrix} \frac{\partial u}{\partial x}(a) & \frac{\partial u}{\partial y}(a) \\ \frac{\partial v}{\partial x}(a) & \frac{\partial v}{\partial y}(a) \end{pmatrix}$ genau dann eine $\mathbb{C}$-lineare Abbildung, wenn die **Cauchy-Riemann Differentialgleichungen**

$$\frac{\partial u}{\partial x}(a) = \frac{\partial v}{\partial y}(a) \tag{15.5.9}$$

$$\frac{\partial u}{\partial y}(a) = -\frac{\partial v}{\partial x}(a) \tag{15.5.10}$$

erfüllt sind. Die mit $f$ zu identifizierende Funktion $\begin{array}{rcl} g : \mathbb{C} & \to & \mathbb{C} \\ x + iy & \mapsto & u(x,y) + iv(x,y) \end{array}$ ist genau dann **komplex differenzierbar** in $a$, wenn $f$ stetig partiell differenzierbar in $a$ ist und die Cauchy-Riemannschen Differentialgleichungen gelten.

**Definition 15.5.17.** *Es sei $U \subseteq \mathbb{C}$ offen, dann heißt eine Funktion $f : U \to \mathbb{C}$* **holomorph**, *wenn $f$ in jedem Punkt $a \in U$ komplex differenzierbar ist. $f : U \to \mathbb{C}$ heißt* **analytisch**, *wenn es für jedes $a \in U$ ein $r > 0$ und eine Potenzreihe $\sum_{n=0}^{\infty} c_n(z-a)^n$ gibt,* *so daß* $\{z \in \mathbb{C} : |z-a| < r\} \subseteq U$ *ist,* $\sum_{n=0}^{\infty} c_n(z-a)^n$ *für* $|z-a| < r$ *konvergiert und* $f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n$ *auf* $\{z \in \mathbb{C} : |z-a| < r\}$ *gilt.*

## 15.6 Höhere Ableitungen und der Satz von Taylor

**Definition 15.6.1.** *Sind* $X_1, ..., X_n, Y$ *Banachräume über demselben Körper* $\mathbb{K}$*, dann ist*

$$L(X_1, ..., X_n; Y) := \{\phi : X_1 \times ... \times X_n \to Y : \phi \textit{ ist multilinear und stetig}\} \tag{15.6.1}$$

*die Menge aller stetigen multilinearen Abbildungen und wir schreiben* $L^n(X;Y)$ *statt* $L(X, ..., X; Y)$*, falls* $X_1 = ... = X_n = X$ *ist.*

**Satz 15.6.2.** *Sind* $X_1, ..., X_n, Y$ *Banachräume über demselben Körper* $\mathbb{K}$*, dann wird* $L(X_1, ..., X_n; Y)$ *durch*

$$\begin{aligned} |||\phi||| \ &:= \ \inf\{C \in ]0,\infty[ : \|\phi[(x_1, ..., x_n)]\| \le C\|x_1\| \cdots \|x_n\| \textit{ für alle} \\ &\qquad\qquad (x_1, ..., x_n) \in X_1 \times ... \times X_n\} \\ &= \ \sup\{\|\phi[(x_1, ..., x_n)]\| : \|x_1\|, ..., \|x_n\| \le 1\} \end{aligned} \tag{15.6.2}$$

*zu einem Banachraum. Es gilt*

$$\|\phi[(x_1, ..., x_n)]\| \le |||\phi|||\, \|x_1\| \cdots \|x_n\| \tag{15.6.3}$$

*für alle* $(x_1, ..., x_n) \in X_1 \times ... \times X_n$*.*

$$\begin{array}{rcl} L(X_1, L(X_2;Y)) & \to & L(X_1, X_2; Y) \\ T & \mapsto & \begin{array}{rcl} \phi_T : X_1 \times X_2 & \to & Y \\ (x_1, x_2) & \mapsto & (T[x_1])[x_2] \end{array} \end{array} \tag{15.6.4}$$

*ist ein isometrischer Isomorphismus und analog sind die Räume* $L(X_1, ..., X_n; Y)$ *und* $L(X_1, L(X_2, ...L(X_n; Y)))$ *isometrisch isomorph.*

*Beweis.* Im Fall $n = 1$ haben wir schon in Satz 13.9.3 gezeigt, daß $(L(X_1, Y), ||| \cdot |||)$ ein Banachraum ist.

Zu $T \in L(X_1, L(X_2, Y))$ definiere $\begin{array}{rcl} \phi_T : X_1 \times X_2 & \to & Y \\ (x_1, x_2) & \mapsto & (T[x_1])[x_2] \end{array}$, dann ist $\phi_T$ offenbar bilinear und $\|(T[x_1])[x_2]\| \le |||T[x_1]||| \cdot \|x_2\| \le |||T||| \cdot \|x_1\| \cdot \|x_2\|$ und damit ist $\phi_T \in L(X_1, X_2; Y)$ und $|||\phi_T||| \le |||T|||$. Ist umgekehrt $\phi \in L(X_1, X_2; Y)$ so definiere $(T_\phi[x_1])[x_2] := \phi[x_1, x_2]$, dann ist $\|(T_\phi[x_1])[x_2]\| \le |||\phi||| \cdot \|x_1\| \cdot \|x_2\|$, daher ist $T_\phi[x_1] : X_2 \to Y \in L(X_2, Y)$ mit $|||T_\phi[x_1]||| \le |||\phi||| \cdot \|x_1\|$ und somit $\begin{array}{rcl} T_\phi : X_1 & \to & L(X_2, Y) \\ x_1 & \mapsto & T_\phi[x_1] \end{array} \in L(X_1, L(X_2, Y))$.

$$\begin{array}{rcl} L(X_1, L(X_2, Y)) & \to & L(X_1, X_2; Y) \\ T & \mapsto & \phi_T \end{array} \quad \text{und} \quad \begin{array}{rcl} L(X_1, X_2; Y) & \to & L(X_1, L(X_2, Y)) \\ \phi & \mapsto & T_\phi \end{array}$$

sind also zueinander inverse stetige lineare Abbildungen und damit $T_{\phi_T} = T$. Damit ist $|||T||| = |||T_{\phi_T}||| \le |||\phi_T||| \le |||T|||$, also $T \mapsto \phi_T$ auch isometrisch. □

**Definition 15.6.3.** *Es seien $X$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen, $a \in U$ und $f : U \to Y$ eine differenzierbare Abbildung. Ist nun*

$$\begin{array}{rcl} Df : U & \to & L(X,Y) \\ x & \mapsto & (Df)(x) \end{array}$$

*im Punkt $a \in U$ differenzierbar, so heißt*

$$(D^2 f)(a) := (D(Df))(a) \in L(X, L(X,Y)) \sim L^2(X;Y)$$

*die zweite Ableitung von $f$ an der Stelle $a$. Dann definiert man induktiv: Ist für $m \in \mathbb{N}$ die $m$-te Ableitung $D^m f : U \to L^m(X,Y)$ definiert, und $D^m f$ an der Stelle $a$ differenzierbar, dann ist $(D^{m+1}f)(a) := (D(D^m f))(a)$ die $m+1$-te Ableitung von $f$ an der Stelle $a$. Existiert die Funktion $D^m f : U \to L^m(X;Y)$ und ist sie stetig, dann heißt $f : U \to Y$ $m$-mal stetig differenzierbar.*

$$\mathcal{C}^m(U,Y) := \{f : U \to Y : f \text{ ist } m\text{-mal stetig differenzierbar}\} \tag{15.6.5}$$

$$\mathcal{C}^\infty(U,Y) := \bigcap_{m\in\mathbb{N}} \mathcal{C}^m(U,Y) \tag{15.6.6}$$

**Lemma 15.6.4.** *Es seien $X, Y$ und $Z$ Banachräume über $\mathbb{K}$, $\Phi : X \times Y \to Z$ sei stetig und bilinear. Dann ist $\Phi \in \mathcal{C}^\infty(X \times Y, Z)$ und $(D^m\Phi)(x,y) = \mathbf{0}$ für alle $m \geq 3$, $x \in X$ und $y \in Y$.*

*Beweis.* Nach Lemma 15.1.13 ist für alle $(a,b) \in X \times Y$ die Ableitung $(D\Phi)(a,b)$ von $\Phi$ im Punkt $(a,b)$ gegeben als

$$\begin{array}{rcll} (D\Phi)(a,b) : X \times Y & \to & Z & \in L(X \times Y, Z) \\ (x,y) & \mapsto & \Phi[a,y] + \Phi[x,b] & \end{array}$$

Da $\Phi$ stetig und bilinear ist, gibt es nach Satz 15.6.2 ein $c \in ]0,\infty[$ mit $\|\Phi[x,y]\| \leq c\|x\| \cdot \|y\|$ für alle $x \in X$ und $y \in Y$. Daher ist

$$|||(D\Phi)(a,b)||| = \sup_{\|(x,y)\|=1} \|(D\Phi)(a,b)[x,y]\| = \sup_{\substack{\|x\|\leq 1\\ \|y\|\leq 1}} \|(D\Phi)(a,b)[x,y]\| \leq c(\|a\| + \|b\|),$$

also $\begin{array}{rcl} D\Phi : X \times Y & \to & L(X \times Y, Z) \\ (a,b) & \mapsto & (D\phi)(a,b) \end{array} \in L(X\times Y, L(X\times Y, Z))$, weshalb $D\Phi$ nach Beispiel 15.1.8 differenzierbar ist und $(D^2\Phi)(a,b) = (D(D\Phi))(a,b) = D\Phi$ konstant ist. Daher existieren alle höheren Ableitungen $D^m\Phi$ und es gilt $D^m\Phi = \mathbf{0}$ für $m \geq 3$. $\square$

**Lemma 15.6.5.** *Es seien $X, Y$ und $Z$ Banachräume über $\mathbb{K}$, $U \subseteq X$ offen, $V \subseteq Y$ offen, $m \in \mathbb{N} \cup \{\infty\}$, $f \in \mathcal{C}^m(U,Y)$ mit $f(U) \subseteq V$, $g \in \mathcal{C}^m(V,Z)$, dann gilt*

$$h := g \circ f \in \mathcal{C}^m(U,Z) \tag{15.6.7}$$

*Beweis.* durch Induktion nach $m$:

- $m = 1$: Aus der Kettenregel folgt:

$$h'(a) = g'(f(a)) \circ f'(a) \tag{15.6.8}$$

und weil die Verknüpfung

$$\begin{array}{rcl} \circ L(X,Y) \times L(Y,Z) & \to & L(X,Z) \\ (u,v) & \mapsto & v \circ u \end{array}$$

stetig und bilinear ist, folgt die Behauptung für $m = 1$ aus Beispiel 15.6.4.

- $m \to m+1$: Nach Voraussetzung sind $f$ und $g$ $m+1-$mal stetig differenzierbar, daher ist nach Induktionsvoraussetzung $g' \circ f$ $m-$mal stetig differenzierbar damit ist nach (15.6.8) auch $h'$ $m-$mal stetig differenzierbar, also $h \in \mathcal{C}^{m+1}(U,Z)$

□

**Definition 15.6.6.** *Es seien $V, W$ Vektorräume, $n \in \mathbb{N}$ mit $n \geq 2$ und $\phi : V^n \to W$ sei multilinear. Dann heißt $\phi$* **symmetrisch***, wenn*

$$\phi(v_{\pi(1)}, ..., v_{\pi(n)}) = \phi(v_1, ..., v_n) \tag{15.6.9}$$

*für jede Permutation $\pi \in \mathcal{S}_n$ und jedes $(v_1, ..., v_n) \in V^n$ gilt.*

**Satz 15.6.7.** *Es seien $X$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ sei offen und $f \in \mathcal{C}^2(U,Y)$, dann gilt:*

$$(D^2 f)(a)[x_1, x_2] = (D^2 f)(a)[x_2, x_1] \tag{15.6.10}$$

*für alle $a \in U$ und $x_1, x_2 \in X$.*

*Beweis.* Zu $a \in U$ sei $s = s(a) > 0$ so gewählt, daß $K(a, 2s) = \{x \in X : \|x-a\| < 2s\} \subseteq U$ ist. Ferner seien $x_1, x_2 \in X$ mit $\|x_1\|, \|x_2\| < s$ und dazu definiere

$$\begin{array}{rcl} g : K(a,s) & \to & Y \\ x & \mapsto & f(x + x_1) - f(x) \end{array}$$

$$\begin{array}{rcl} r : K(\mathbf{0}, s) \times K(\mathbf{0}, s) & \to & Y \\ (x_1, x_2) & \mapsto & \displaystyle\int_0^1 \int_0^1 \big((D^2 f)(a + t_1 x_1 + t_2 x_2)[x_1] - (D^2 f)(a)[x_1]\big)\, dt_1 [x_2] dt_2 \end{array}$$

Dann ist $f(a + x_1 + x_2) - f(a + x_2) - f(a + x_1) + f(a) = g(a + x_2) - g(a)$ und da $a + t_1 x_1 + t_2 x_2 \in K(a, 2s) \subseteq U$ für $t_1, t_2 \in [0,1]$ ist, folgt aus dem Mittelwertsatz:

$$\begin{aligned} f(a + x_1 + x_2) - f(a + x_2) - f(a + x_1) + f(a) &= g(a + x_2) - g(a) \qquad (15.6.11) \\ = \int_0^1 (Dg)(a + t_2 x_2)[x_2] dt_2 &= \int_0^1 ((Df)(a + x_1 + t_2 x_2)[x_2] - (Df)(a + t_2 x_2)[x_2]) dt_2 \\ = \int_0^1 \left( \int_0^1 (D^2 f)(a + t_1 x_1 + t_2 x_2)[x_1] dt_1 \right) [x_2] dt_2 &= (D^2 f)(a)[x_1, x_2] + r(x_1, x_2) \end{aligned}$$

Für

$$\begin{array}{rcl}\tilde{g} : K(a,s) & \to & Y \\ x & \mapsto & f(x+x_2) - f(x)\end{array}$$

$$\begin{array}{rcl}\tilde{r} : K(\mathbf{0},s) \times K(\mathbf{0},s) & \to & Y \\ (x_1,x_2) & \mapsto & \int\limits_0^1 \int\limits_0^1 \big((D^2 f)(a+t_1x_2+t_2x_1)[x_2] - (D^2 f)(a)[x_2]\big)\, dt_1[x_1]dt_2\end{array}$$

und analog zu (15.6.11) folgt:

$$\begin{aligned} f(a+x_1+x_2) - f(a+x_1) - f(a+x_2) + f(a) &= \tilde{g}(a+x_1) - \tilde{g}(a) \qquad (15.6.12)\\ = \int\limits_0^1 (D\tilde{g})(a+t_2x_1)[x_1]dt_2 &= \int\limits_0^1 ((Df)(a+t_2x_1+x_2)[x_1] - (Df)(a+t_2x_1)[x_1])dt_2 \\ = \int\limits_0^1 \left( \int\limits_0^1 (D^2 f)(a+t_2x_1+t_1x_2)[x_2]dt_1 \right) [x_1]dt_2 &= (D^2 f)(a)[x_2,x_1] + \tilde{r}(x_1,x_2)\end{aligned}$$

Damit ist $(D^2 f)(a)[x_1,x_2] - (D^2 f)(a)[x_2,x_1] = \tilde{r}(x_1,x_2) - r(x_1,x_2)$ und wegen

$$\begin{aligned} &\max\{\|\tilde{r}(x_1,x_2)\|, \|r(x_1,x_2)\|\} \\ &\quad\leq \sup\{|||(D^2 f)(a+t_1x_1+t_2x_2) - (D^2 f)(a)||| \cdot \|x_1\| \cdot \|x_2\| : t_1,t_2 \in [0,1]\}\end{aligned}$$

ist

$$\begin{aligned} &\|(D^2 f)(a)[x_1,x_2] - (D^2 f)(a)[x_2,x_1]\| \\ &\quad\leq 2\sup\{|||(D^2 f)(a+t_1x_1+t_2x_2) - (D^2 f)(a)||| \cdot \|x_1\| \cdot \|x_2\| : t_1,t_2 \in [0,1]\}\end{aligned}$$

Für jedes $\varepsilon > 0$ gibt es, da $D^2 f$ stetig ist, ein $r_\varepsilon \in ]0,s[$ mit

$$\sup\left\{|||(D^2 f)(a+t_1x_1+t_2x_2) - (D^2 f)(a)||| \cdot \|x_1\| \cdot \|x_2\| : t_1,t_2 \in [0,1], \|x_1\|, \|x_2\| < r_\varepsilon\right\} \leq 2\varepsilon,$$

daher ist

$$\|(D^2 f)(a)[x_1,x_2] - (D^2 f)(a)[x_2,x_1]\| \leq \varepsilon \|x_1\| \cdot \|x_2\| \qquad (15.6.13)$$

für alle $x_1, x_2 \in X$ mit $\|x_1\|, \|x_2\| < r_\varepsilon$. Wegen Bilinearität gilt (15.6.13) sogar für alle $x_1, x_2 \in X$ und da dann (15.6.13) für alle $\varepsilon > 0$ erfüllt ist, folgt $(D^2 f)(a)[x_1,x_2] = (D^2 f)(a)[x_2,x_1]$ für alle $x_1, x_2 \in X$. □

**Korollar 15.6.8.** *Es seien $X$ und $Y$ Banachräume über demselben Körper $\mathbb{K}$, $U \subseteq X$ offen und $f : U \to Y$ sei $m$-mal stetig differenzierbar (für ein $m \geq 2$) und $a \in U$. Dann ist $(D^m f)(a) \in L^m(X,Y)$ eine symmetrische Abbildung.*

*Beweis.* durch Induktion nach $m$:

- $m = 2$ ist Satz 15.6.7

- $m \to m+1$: Wegen $(D^{m+1}f)(a) = D^2(D^{m-1}f)(a)$ folgt für alle $x_1, ..., x_m, x_{m+1} \in X$ aus Satz 15.6.7:

$$\begin{aligned}&(D^{m+1}f)(a)[x_1, ..., x_{m-1}, x_m, x_{m+1}] = D^2((D^{m-1}f)(a)[x_1, ..., x_{m-1}])[x_m, x_{m+1}] \\ =\ & D^2((D^{m-1}f)(a)[x_1, ..., x_{m-1}])[x_{m+1}, x_m] = (D^{m+1}f)(a)[x_1, ..., x_{m-1}, x_{m+1}, x_m]\end{aligned}$$

  Da nach Induktionsvoraussetzung $(D^m f)(a)[x_1, ..., x_m] = (D^m f)(a)[x_{\pi(1)}, ..., x_{\pi(m)}]$ für alle $\pi \in \mathcal{S}_m$ erfüllt ist und sich jedes $\sigma \in \mathcal{S}_{m+1}$ als Komposition von Transpositionen, also als Kompostition der Transposition $\tau = \begin{pmatrix} 1 & \cdots & m-1 & m & m+1 \\ 1 & \cdots & m-1 & m+1 & m \end{pmatrix}$ und der Permutation $\begin{pmatrix} 1 & \cdots & m & m+1 \\ \pi(1) & \cdots & \pi(m) & m+1 \end{pmatrix}$ mit $\pi = \pi(\sigma) \in \mathcal{S}_m$ schreiben läßt, folgt die Behauptung. □

**Definition 15.6.9.** *Es sei $Y$ ein $\mathbb{K}-$Banachraum, $m, d \in \mathbb{N}$ mit $d \geq 2$ und $U \subseteq \mathbb{K}^d$ offen. Eine Abbildung $f : U \to Y$ heißt $m-$***mal (stetig) partiell differenzierbar** *, wenn für jede Wahl von $m$ Indizes $j_1, ..., j_m \in \{1, ..., d\}$ die partiellen Ableitungen*

$$\frac{\partial^m f}{\partial x_{j_m} \cdots \partial x_{j_1}} := \frac{\partial}{\partial x_{j_m}} \left( \cdots \left( \frac{\partial f}{\partial x_{j_1}} \right) \cdots \right)$$

*existieren (und stetig sind).*

**Satz 15.6.10.** *Es sei $Y$ ein $\mathbb{K}-$Banachraum; $m, d \in \mathbb{N}$ mit $d \geq 2$, $U \subseteq \mathbb{K}^d$ offen und $f : U \to Y$, dann sind äquivalent:*

*a)* $f \in \mathcal{C}^m(U, Y)$

*b)* *$f$ ist $m-$mal stetig partiell differenzierbar.*

*In diesem Fall ist für jedes $q \in \mathbb{N}$ mit $q \leq m$, für jede Wahl $j_1, ..., j_q \in \{1, ..., d\}$ und jedes $\pi \in \mathcal{S}_q$:*

$$\frac{\partial^q f}{\partial x_{j_1} \cdots \partial x_{j_q}} = \frac{\partial^q f}{\partial x_{\pi(j_1)} \cdots \partial x_{\pi(j_q)}} \tag{15.6.14}$$

*und wenn $h_l = \sum_{j_l=1}^{d} h_l^{j_l} \underline{e}_{j_l}$ ist, so gilt*

$$(D^q f)(a)[h_1, ..., h_q] = \sum_{j_1, ..., j_q = 1}^{d} \frac{\partial}{\partial x_{j_q}} \left( \cdots \left( \frac{\partial f}{\partial x_{j_1}} \right) \cdots \right)(a) h_1^{j_1} \cdots h_q^{j_q}. \tag{15.6.15}$$

*Beweis.* Es sei $f$ $q-$mal stetig differenzierbar mit $q \leq m$ und $h_1, ..., h_q \in \mathbb{K}^d$, dann folgt induktiv aus Lemma 15.5.3

$$\begin{aligned}&(D^q f)(a)[h_1, ..., h_q] = D(\cdots D((Df)(a)[h_1])[h_2] \cdots)[h_q] \\ =\ & D(\cdots D((D_{h_1} f)(a))[h_2] \cdots)[h_q] = D_{h_q} \cdots D_{h_2} D_{h_1} f(a)\end{aligned}$$

insbesondere folgt für die Standardbasisvektoren $\underline{e}_1, ..., \underline{e}_d$ in $\mathbb{K}^d$:

$$(D^q f)(a)[\underline{e}_{j_1}, ..., \underline{e}_{j_q}] = \frac{\partial}{\partial x_{j_q}} \left( \cdots \left( \frac{\partial f}{\partial x_{j_1}} \right) \cdots \right)(a) \tag{15.6.16}$$

und wenn $h_l = \sum\limits_{j_l=1}^{d} h_l^{j_l} \underline{e}_{j_l}$ ist, so folgt aus $(Df)(a) \in L(\mathbb{K}^d, Y)$ induktiv:

$$(D^q f)(a)[h_1, ..., h_q] = \sum_{j_1,...,j_q=1}^{d} \frac{\partial}{\partial x_{j_q}} \left( \cdots \left( \frac{\partial f}{\partial x_{j_1}} \right) \cdots \right)(a) h_1^{j_1} \cdots h_q^{j_q} \tag{15.6.17}$$

a)⇒b) Ist $f \in \mathcal{C}^m(U, Y)$, dann definiert für jedes $q \leq m$ die linke Seite von (15.6.16) eine stetige Funktion $\begin{array}{rcl} (D^q f)(\cdot)[\underline{e}_{j_1}, ..., \underline{e}_{j_q}] : U & \to & Y \\ a & \mapsto & (D^q f)(a)[\underline{e}_{j_1}, ..., \underline{e}_{j_q}] \end{array}$ und damit ist nach (15.6.16) auch $\begin{array}{rcl} \frac{\partial^q f}{\partial x_{j_1} \cdots \partial x_{j_q}} : U & \to & Y \\ a & \mapsto & \frac{\partial}{\partial x_{j_q}} \left( \cdots \left( \frac{\partial f}{\partial x_{j_1}} \right) \cdots \right)(a) \end{array}$ stetig für jede Wahl von $q$ Indizes $i_1, ..., i_q \in \{1, ..., d\}$, dh. $f$ ist $m-$mal stetig partiell differenzierbar.

b)⇒a) durch Induktion nach $m$:

- $m = 1$: Nach Satz 15.5.9 ist $f$ genau dann stetig differenzierbar, wenn $f$ stetig partiell differenzierbar ist.
- $m \to m + 1$: Es sei $f$ $m + 1-$mal stetig partiell differenzierbar und nach Induktionsvoraussetzung ist für alle $h_l = \sum\limits_{j_l=1}^{d} h_l^{j_l} \underline{e}_{j_l} \in \mathbb{K}^d$, $l = 1, ..., m$ wegen

  $$(D^m f)(a)[h_1, ..., h_m] = \sum_{j_1,...,j_m=1}^{d} \frac{\partial}{\partial x_{j_m}} \cdots \frac{\partial f}{\partial x_{j_1}}(a) h_1^{j_1} \cdots h_m^{j_m}$$

  die Funktion

  $$\begin{array}{rcl} (D^m f)(\cdot)[h_1, ..., h_m] : U & \to & Y \\ a & \mapsto & (D^m f)(a)[h_1, ..., h_m] \end{array}$$

  stetig partiell differenzierbar und daher existiert nach Satz 15.5.9 die Funktion $D(D^m f(\cdot)[h_1, ..., h_m])$ und ist stetig. (15.6.15) gilt dann auch für $q = m + 1$

  $$(D^{m+1} f)(a)[h_1, ..., h_{m+1}] = \sum_{j_1,...,j_{m+1}=1}^{d} \frac{\partial}{\partial x_{j_{m+1}}} \cdots \frac{\partial f}{\partial x_{j_1}}(a) h_1^{j_1} \cdots h_{m+1}^{j_{m+1}} \tag{15.6.18}$$

  und da nach Voraussetzung für jede Wahl von Indizes $j_1, ..., j_{m+1} \in \{1, ..., d\}$ die partielle Ableitung

  $$\begin{array}{rcl} \frac{\partial}{\partial x_{j_{m+1}}} \cdots \frac{\partial f}{\partial x_{j_1}} : U & \to & Y \\ a & \mapsto & \frac{\partial}{\partial x_{j_{m+1}}} \cdots \frac{\partial f}{\partial x_{j_1}}(a) \end{array}$$

  stetig ist, so zeigt (15.6.18), daß $(D^{m+1} f)(a) \in L^{m+1}(\mathbb{R}^d, Y)$ ist.

Die Gleichheit (15.6.14) folgt in diesem Fall aus (15.6.16) und Korollar 15.6.8, denn da $(D^q f)(a)$ symmetrisch ist, gilt

$$(D^q f)(a)[h_1, ..., h_q] = (D^q f)(a)[h_{\pi(1)}, ..., h_{\pi(q)}]$$

für jedes $q \leq m$, $h_1, ..., h_q \in \mathbb{K}^d$ und $\pi \in \mathcal{S}_q$. $\square$

**Korollar 15.6.11** (Schwarz)**.** *Es sei $U \subseteq \mathbb{K}^d$ offen und $f : U \to \mathbb{K}$, dann sind äquivalent:*

*a)* $f \in \mathcal{C}^2(U, \mathbb{K})$

*b)* $f, \dfrac{\partial f}{\partial x_j}, \dfrac{\partial^2 f}{\partial x_j \partial x_k} \in \mathcal{C}(U, \mathbb{K})$ *für alle* $j, k \in \{1, ..., d\}$.

*In diesem Fall gilt:* $\dfrac{\partial^2 f}{\partial x_j \partial x_k} = \dfrac{\partial^2 f}{\partial x_k \partial x_j}$ *für alle* $j, k \in \{1, ..., d\}$.

**Bemerkung 15.6.12.** Ist $f$ nur zweimal partiell differenzierbar (aber eine zweite partielle Ableitung nicht stetig), so wird die Folgerung von Korollar 15.6.11 falsch. Beispiel:

$$\begin{array}{rcl} f : \mathbb{R}^2 & \to & \mathbb{R} \\ (x, y) & \mapsto & \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{für } (x,y) \neq (0,0) \\ 0 & \text{für } (x,y) = (0,0) \end{cases} \end{array}$$

**Satz 15.6.13.** *[Taylor] Es seien $X$ und $Y$ $\mathbb{K}-$Banachräume, $U \subseteq X$ offen, $a \in U$, $m \in \mathbb{N}$ und $f \in \mathcal{C}^m(U, Y)$. Für $h \in X$ sei*

$$(D^k f)(a)[h]^k := \begin{cases} f(a) & \text{für } k = 0 \\ (D^k f)(a) \underbrace{[h, ..., h]}_{k-mal} & \text{für } k \in \{1, ..., m\} \end{cases} \tag{15.6.19}$$

*dann gilt im Fall von $[\![a, a+h]\!] \subseteq U$*

$$f(a+h) = \sum_{k=0}^{m} \frac{1}{k!} (D^k f)(a)[h]^k + R(f, m, a, h) \tag{15.6.20}$$

*mit*

$$R(f, m, a, h) := \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} \Big( (D^m f)(a+th)[h]^m - (D^m f)(a)[h]^m \Big) dt. \tag{15.6.21}$$

*Beweis.* durch Induktion nach $m$:

- $m = 1$: Nach dem Mittelwertsatz in Integralform gilt:

$$\begin{aligned} f(a+h) &= f(a) + \int_0^1 (Df)(a+th)[h] dt \\ &= f(a) + (Df)(a)[h] + \int_0^1 ((Df)(a+th)[h] - (Df)(a)[h]) dt \end{aligned}$$

- $m \to m+1$: Nach Induktionsvoraussetzung und wegen $\int\limits_0^1 \frac{(1-t)^{m-1}}{(m-1)!}dt = \frac{1}{m!}$ ist

$$
\begin{aligned}
f(a+h) &= \sum_{k=0}^{m} \frac{1}{k!}(D^k f)(a)[h]^k \\
&\quad + \int\limits_0^1 \frac{(1-t)^{m-1}}{(m-1)!}\Big((D^m f)(a+th)[h]^m - (D^m f)(a)[h]^m\Big)dt \\
&= \sum_{k=0}^{m-1} \frac{1}{k!}(D^k f)(a)[h]^k + \int\limits_0^1 \frac{(1-t)^{m-1}}{(m-1)!}(D^m f)(a+th)[h]^m dt
\end{aligned}
$$

Partielle Integration des letzten Integrals mit $u_m(t) := (D^m f)(a+th)[h]^m$ und $v_m'(t) := \frac{(1-t)^{m-1}}{(m-1)!}$ mit $v_m(t) = -\frac{[1-t)^m}{m!}$ und $u_m'(t) = (D^{m+1}f)(a+th)[h]^{m+1}$ ergibt:

$$
\begin{aligned}
f(a+h) &= \sum_{k=0}^{m-1} \frac{1}{k!}(D^k f)(a)[h]^k - \Big((D^m f)(a+th)[h]^m \frac{(1-t)^m}{m!}\Big)\Big|_{t=0}^{t=1} \\
&\quad + \int\limits_0^1 \frac{(1-t)^m}{m!}(D^{m+1} f)(a+th)[h]^{m+1} dt \\
&= \sum_{k=0}^{m} \frac{1}{k!}(D^k f)(a)[h]^k + \int\limits_0^1 \frac{(1-t)^m}{m!}(D^{m+1} f)(a+th)[h]^{m+1} dt \\
&= \sum_{k=0}^{m+1} \frac{1}{k!}(D^k f)(a)[h]^k \\
&\quad + \int\limits_0^1 \frac{(1-t)^m}{m!}\Big((D^{m+1} f)(a+th)[h]^{m+1} - (D^{m+1} f)(a)[h]^{m+1}\Big)dt
\end{aligned}
$$

wobei die letzte Gleichheit wieder aus $\int\limits_0^1 \frac{(1-t)^m}{m!}dt = \frac{1}{(m+1)!}$ folgt. □

**Bemerkung 15.6.14.** Es seien $X$ und $Y$ Banachräume, $U \subseteq X$ offen, $a \in U$, $m \in \mathbb{N}$ und $f \in \mathcal{C}^m(U,Y)$. Für das Restglied $R(f,m,a,h)$ im Satz von Taylor gilt:

$$
\limsup_{\|h\| \to 0} \frac{\|R(f,m,a,h)\|(m-1)!}{\|h\|^m} = 0 \tag{15.6.22}
$$

*Beweis.* Da $U$ offen und $a \in U$ ist, gibt es ein $\varepsilon = \varepsilon(a) > 0$ mit $K(a, \varepsilon) \subseteq U$ und daher ist $[\![a, a+h]\!] \subseteq K(a, \varepsilon) \subseteq U$ für jedes $h \in X$ mit $\|h\| < \varepsilon$. Die Gleichheit

$$f(a+h) = \sum_{k=0}^{m} \frac{1}{k!}(D^k f)(a)[h]^k + R(f, m, a, h)$$

mit

$$R(f, m, a, h) := \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!}\Big((D^m f)(a+th)[h]^m - (D^m f)(a)[h]^m\Big)dt$$

aus dem Satz von Taylor gilt also für alle $h \in X$ mit $\|h\| < \varepsilon$. Es ist

$$\begin{aligned}\|R(f, m, a, h)\| &= \left\| \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!}\Big((D^m f)(a+th)[h]^m - (D^m f)(a)[h]^m\Big)dt \right\| \\ &\le \frac{1}{(m-1)!}\int_0^1 |||(D^m f)(a+th) - (D^m f)(a)||| \cdot ||h||^m dt,\end{aligned}$$

denn für jedes $t \in [0,1]$ ist $0 \le (1-t)^{m-1} \le 1$. Da $f \in \mathcal{C}^m(U, Y)$ vorausgesetzt ist, ist $\begin{array}{rcl} D^m f : U & \to & L^m(X, Y) \\ x & \mapsto & (D^m f)(x) \end{array}$ stetig, also $\lim\limits_{h \to \mathbf{0}} |||(D^m f)(a+th) - (D^m f)(a)||| = 0$. Ferner läßt sich für jedes $\varepsilon > 0$ ein $\delta = \delta(\varepsilon, a) > 0$ mit $K(a, \delta) \subseteq U$ und $|||(D^m f)(a+\xi) - (D^m f)(a)||| < \varepsilon$ für alle $\xi \in X$, $\|\xi\| < \delta$. Da $\mathbf{1}_{[0,1]}\varepsilon$ daher eine $\widehat{\lambda}-$integrierbare Majorante für $t \in [0,1]$, $\|h\| < \delta$ ist, gilt $\int_0^1 |||(D^m f)(a+th) - (D^m f)(a)|||dt < \varepsilon$. □

## 15.7 Lokale Extrema

**Definition 15.7.1.** *Es sei $d \in \mathbb{N}$, $d \ge 2$, $U \subseteq \mathbb{R}^d$ offen, $a \in U$ und $f : U \to \mathbb{R}$ sei zweimal stetig differenzierbar. Dann heißt die Matrix*

$$H_f(a) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_d}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_d}(a) \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_d \partial x_1}(a) & \frac{\partial^2 f}{\partial x_d \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_d^2}(a) \end{pmatrix} \tag{15.7.1}$$

*die* **Hessematrix** *von $f$ in $a$.*

**Bemerkung 15.7.2.** Ist $d \in \mathbb{N}$, $d \ge 2$, $U \subseteq \mathbb{R}^d$ offen, $f \in \mathcal{C}^2(U, \mathbb{R})$ und $a \in U$, dann ist:

- Die Hessematrix die darstellende Matrix der Bilinearform $(D^2f)(a)$ bezüglich der Standardbasis, dh. sind $\begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}$ und $\begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix}$ die Koordinatenvektoren von $\underline{x}$ bzw. $\underline{y}$ in der Standardbasis, dann ist

$$(D^2f)(a)[\underline{x}, \underline{y}] = \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}, H_f(a) \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix} \right\rangle =: \langle \underline{x}, H_f(a)\underline{x} \rangle.$$

Einsetzen der Definition des Gradienten in Satz 15.6.13 ergibt

$$f(a+h) = f(a) + \langle (\mathrm{grad} f)(a), h \rangle + \frac{1}{2}\langle h, H_f(a)h \rangle + R(f, 2, a, h) \tag{15.7.2}$$

mit

$$\limsup_{\substack{|h| \to 0, \\ h \neq 0}} \frac{|R(f, 2, a, h)|}{h^2} = 0$$

- Nach Korollar 15.6.11 ist $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$, also ist die Matrix $H_f(a)$ selbstadjungiert. Nach Lemma 12.2.4 sind alle Eigenwerte $\lambda_1, ..., \lambda_d$ von $H_f(a)$ reell und es gibt eine Orthonormalbasis $\underline{v}_1, ..., \underline{v}_d$ von $\mathbb{R}^d$ aus Eigenvektoren von $H_f(a)$.

**Satz 15.7.3.** *Es sei $d \in \mathbb{N}$, $U \subseteq \mathbb{R}^d$ offen, $f : U \to \mathbb{R}$ sei zweimal stetig differenzierbar und $a \in U$ mit $(\mathrm{grad} f)(a) = \mathbf{0}$. Es sei $H_f(a)$ die Hessematrix von $f$ in $a$ und $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_d$ die Eigenwerte von $H_f(a)$. Dann gilt:*

*a) Ist $0 < \lambda_1$, so hat $f$ in $a$ ein isoliertes lokales Minimum.*

*b) Ist $\lambda_d < 0$, so hat $f$ in $a$ ein isoliertes lokales Maximum.*

*c) Ist $\lambda_1 < 0 < \lambda_d$, so hat $f$ in $a$ kein lokales Extremum.*

*Beweis.* Da $a \in U$ und $U$ offen ist, gibt es $r > 0$ mit $a + h \in U$ für alle $h \in \mathbb{R}^d$ mit $\|h\| < r$. Da $(\mathrm{grad} f)(a) = \mathbf{0}$ ist, folgt

$$f(a+h) = f(a) + \frac{1}{2}\langle h, H_f(a)h \rangle + R(f, a, 2, h) \tag{15.7.3}$$

mit

$$\limsup_{\substack{|h| \to 0, \\ h \neq 0}} \frac{|R(f, 2, a, h)|}{h^2} = 0 \tag{15.7.4}$$

aus (15.7.2).

a) Ist $0 < \lambda_1$, dann ist $\langle H_f(a)\underline{x}, \underline{x}\rangle \geq \lambda_1 \|\underline{x}\|^2$ für alle $\underline{x} \in \mathbb{R}^d$, denn für die selbstadjungierte Matrix $H_f(a)$ gibt es eine Orthonormalbasis $(\underline{e}_1, ..., \underline{e}_d)$ von $\mathbb{R}^d$ aus Eigenvektoren von $H_f(a)$ und zu $\underline{x} = \sum_{j=1}^{d} x_j \underline{e}_j$ ist

$$\langle H_f(a)\underline{x}, \underline{x}\rangle = \sum_{j,k=1}^{d} x_k x_j \langle \lambda_k \underline{e}_k, \underline{e}_j\rangle = \sum_{j=1}^{d} x_j^2 \lambda_j \geq \lambda_1 \|\underline{x}\|^2.$$

Wegen (15.7.4) gibt es ein $\delta > 0$ mit $|R(f, a, 2, h)| \leq \frac{\lambda_1}{4}\|h\|^2$ für alle $h \in \mathbb{K}^d$ mit $\|h\| < \delta$ und daher ist

$$\begin{aligned} f(a+h) &= f(a) + \frac{1}{2}\langle h, H_f(a)h\rangle + R(f, 2, a, h) \geq f(a) + \frac{\lambda_1}{2}\|h\|^2 - \frac{\lambda_1}{4}\|h\|^2 \\ &= f(a) + \frac{\lambda_1}{4}\|h\|^2 \end{aligned}$$

für jedes $h \in \mathbb{R}^d$ mit $\|h\| < \delta$, dh. in $a$ ist ein lokales Minimum von $f$.

b) Folgt durch Anwenden von Teil a) auf $-f$.

c) Ist $\lambda_1 < 0$, $\lambda_d > 0$ und $H_f(a)\underline{v}_1 = \lambda_1 \underline{v}_1$, $H_f(a)\underline{v}_d = \lambda_d \underline{v}_d$, dann ist

$$\lambda_1\|\underline{v}_1\|^2 = \langle \underline{v}_1, H_f(a)\underline{v}_1\rangle < 0 < \langle \underline{v}_d, H_f(a)\underline{v}_d\rangle = \lambda_d\|\underline{v}_d\|^2$$

Wegen (15.7.4) gibt es $\delta > 0$ mit

$$|R(f, 2, a, t\underline{v}_1)| \leq \frac{t^2\lambda_1}{4}\|\underline{v}_1\|^2 \quad \text{und} \quad |R(f, 2, a, t\underline{v}_d)| \leq \frac{t^2\lambda_d}{4}\|\underline{v}_d\|^2$$

für alle $t \in ]-\delta, \delta[$. Damit ist

$$\begin{aligned} f(a+t\underline{v}_d) &\geq f(a) + t^2\langle \underline{v}_d, H_f(a)\underline{v}_d\rangle - |R(f, 2, a, t\underline{v}_d)| \geq f(a) + \frac{3t^2\lambda_d}{4}\|\underline{v}_d\|^2 > f(a) \\ f(a+t\underline{v}_1) &\leq f(a) + t^2\langle \underline{v}_1, H_f(a)\underline{v}_1\rangle - |R(f, 2, a, t\underline{v}_1)| \leq f(a) + \frac{3t^2\lambda_1}{4}\|\underline{v}_d\|^2 < f(a) \end{aligned}$$

für alle $t \in ]-\delta, \delta[$, dh. $a$ ist kein lokaler Extremalpunkt von $f$. □

**Bemerkung 15.7.4.** Das Kriterium in Satz 15.7.3 deckt nicht alle Fälle ab, zB. $\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^3 \end{array}$ oder $\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^4 \end{array}$ oder $\begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x^2 \end{array}$ fallen aus dem Schema.

**Satz 15.7.5.** *Es seien $n, l \in \mathbb{N}$ mit $l < n$ und $W \subseteq \mathbb{R}^n$ offen, ferner $F : W \to \mathbb{R}$, $h_1 : W \to \mathbb{R}$,..., $h_l : W \to \mathbb{R}$ seien stetig differenzierbar. Für*

$$h = \begin{pmatrix} h_1 \\ \vdots \\ h_l \end{pmatrix} : \begin{array}{ccc} W & \to & \mathbb{R}^l \\ \underline{x} & \mapsto & \begin{pmatrix} h_1(\underline{x}) \\ \vdots \\ h_l(\underline{x}) \end{pmatrix} \end{array}$$

*sei* $M := h^{-1}(\{\underline{0}\}) \neq \emptyset$ *und für alle* $\underline{x} \in h^{-1}(\{\underline{0}\}) = M$ *sei* $(Dh)(\underline{x}) : \mathbb{R}^n \to \mathbb{R}^l$ *surjektiv. Besitzt* $F|_M$ *in* $\underline{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_l \end{pmatrix} \in M$ *ein lokales Extremum („* **Extremum unter den Nebenbedingungen** $h_1(\underline{p}) = 0,\ldots,h_l(\underline{p}) = 0$*“), so gibt es eindeutig bestimmte* $\lambda_1, \ldots, \lambda_l \in \mathbb{R}$ *(„***Lagrangesche Multiplikatoren***“), so daß*

$$\operatorname{grad}(F - \lambda_1 h_1 - \ldots - \lambda_l h_l)(\underline{p}) = \underline{0} \tag{15.7.5}$$

*ist.*

**Rezept 15.7.6.** Es seien $n, l \in \mathbb{N}$ mit $l < n$ und $W \subseteq \mathbb{R}^n$ offen, ferner $F : W \to \mathbb{R}$, $h_1 : W \to \mathbb{R}$,..., $h_l : W \to \mathbb{R}$ seien stetig differenzierbar. Für

$$h = \begin{pmatrix} h_1 \\ \vdots \\ h_l \end{pmatrix} : \begin{array}{rcl} W & \to & \mathbb{R}^l \\ \underline{x} & \mapsto & \begin{pmatrix} h_1(\underline{x}) \\ \vdots \\ h_l(\underline{x}) \end{pmatrix} \end{array}$$

sei $M := h^{-1}(\{\underline{0}\}) \neq \emptyset$ und für alle $\underline{x} \in h^{-1}(\{\underline{0}\}) = M$ sei $(Dh)(\underline{x}) : \mathbb{R}^n \to \mathbb{R}^l$ surjektiv. Um ein lokales Extremum $\underline{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_l \end{pmatrix} \in M$ von $F|_M$ zu bestimmen, löse das Gleichungssystem

$$\begin{array}{rcl} h_1(p_1, \ldots, p_n) & = & 0 \\ & \vdots & \\ h_l(p_1, \ldots, p_n) & = & 0 \\ \partial_1(F - \lambda_1 h_1 - \ldots - \lambda_l h_l)(p_1, \ldots, p_n) & = & 0 \\ & \vdots & \\ \partial_d(F - \lambda_1 h_1 - \ldots - \lambda_l h_l)(p_1, \ldots, p_n) & = & 0 \end{array}$$

nach den Unbekannten $p_1, \ldots, p_n, \lambda_1, \ldots, \lambda_l \in \mathbb{R}$ und teste, ob in $\underline{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_l \end{pmatrix}$ ein lokales Extremum von $F|_M$ vorliegt.

# Kapitel 16

# Integration

## 16.1 Produktmaße und der Satz von Fubini

**Definition 16.1.1.** *Es sei $n \in \mathbb{N}$ und $(X_1, \mathcal{A}_1)$,...,$(X_n, \mathcal{A}_n)$ seien Meßräume, dann heißt*

$$\mathcal{Z} := \{A_1 \times ... \times A_n : A_1 \in \mathcal{A}_1, ..., A_n \in \mathcal{A}_n\}$$

*die Menge aller* **Zylindermengen** *und*

$$\sigma(\mathcal{Z}) =: \mathcal{A}_1 \otimes ... \otimes \mathcal{A}_n \tag{16.1.1}$$

*die* **Produkt $\sigma$-Algebra** *auf $X := X_1 \times ... \times X_n$.*

**Lemma 16.1.2.** *Es seien $(X_1, \mathcal{A}_1)$,...,$(X_n, \mathcal{A}_n)$ Meßräume und $\mathcal{E}_1 \subseteq \mathcal{P}(X_1)$,..., $\mathcal{E}_n \subseteq \mathcal{P}(X_n)$ mit $\sigma(\mathcal{E}_1) = \mathcal{A}_1$,...,$\sigma(\mathcal{E}_n) = \mathcal{A}_n$. Dann gilt*

$$\mathcal{A}_1 \otimes ... \otimes \mathcal{A}_n = \sigma(\mathcal{E}) \tag{16.1.2}$$

*für*

$$\mathcal{E} := \{E_1 \times ... \times E_n : E_1 \in \mathcal{E}_1, ..., E_n \in \mathcal{E}_n\}.$$

*Beweis.* Nach Definition der Zylindermengen $\mathcal{Z} := \{A_1 \times ... \times A_n : A_1 \in \mathcal{A}_1, ..., A_n \in \mathcal{A}_n\}$ ist $\mathcal{E} \subseteq \mathcal{Z}$, also $\sigma(\mathcal{E}) \subseteq \sigma(\mathcal{Z}) = \mathcal{A}_1 \otimes ... \otimes \mathcal{A}_n$. Für $A_1 \in \sigma(\mathcal{E}_1) = \mathcal{A}_1$ ist

$$A_1 \times E_2 \times ... \times E_n \in \sigma(\{E_1 \times E_2 \times ... \times E_n : E_1 \in \mathcal{E}_1\}) \subseteq \sigma(\mathcal{E})$$

bei fixiertem $E_2 \in \mathcal{E}_2$,...,$E_n \in \mathcal{E}_n$ und dann bekommt man induktiv $A_1 \times ... \times A_n \in \sigma(\mathcal{E})$ für jedes $A_1 \in \mathcal{A}_1$,... $A_n \in \mathcal{A}_n$, dh. $\mathcal{Z} \subseteq \sigma(\mathcal{E})$ und damit $\mathcal{A}_1 \otimes ... \otimes \mathcal{A}_n = \sigma(\mathcal{Z}) \subseteq \sigma(\mathcal{E})$. □

**Lemma 16.1.3.** *Es sei $d \in \mathbb{N}$, dann ist*

$$\mathcal{B}(\mathbb{R}^d) = \underbrace{\mathcal{B}(\mathbb{R}) \otimes ... \otimes \mathcal{B}(\mathbb{R})}_{d-mal} \tag{16.1.3}$$

*Beweis.* Nach Satz 13.7.7 sind alle Normen auf $\mathbb{R}^d$ äquivalent und erzeugen daher dieselbe Topologie. Da für $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \in \mathbb{R}^d$ und $\varepsilon > 0$ der offene Würfel $]x_1-\varepsilon, x_1+\varepsilon[\times ... \times]x_d-\varepsilon, x_d+\varepsilon[= \{\underline{y} \in \mathbb{R}^d : \|\underline{x}-\underline{y}\|_\infty < \varepsilon\}$ gerade die offene $\varepsilon$-Kugel um $\underline{x}$ in der $\|\cdot\|_\infty$-Norm ist und da $\mathbb{Q}^d$ dicht ist, liefert

$$\mathcal{E} := \{]x_1 - \frac{1}{n}, x_1 + \frac{1}{n}[\times ... \times]x_d - \frac{1}{n}, x_d + \frac{1}{n}[ : (x_1, ..., x_d) \in \mathbb{Q}^d, n \in \mathbb{N}\}$$

eine abzählbare Basis der Topologie $\mathcal{O}_{\|\cdot\|_\infty} = \mathcal{O}_{\mathbb{R}^d}$. Offenbar sind alle Elemente von $\mathcal{E}$ auch Zylindermengen von $\mathcal{B}(\mathbb{R}) \otimes ... \otimes \mathcal{B}(\mathbb{R})$ und daher nach Lemma 14.1.8

$$\mathcal{B}(\mathbb{R}^d) = \sigma(\mathcal{E}) \subseteq \sigma(\mathcal{Z}) = \underbrace{\mathcal{B}(\mathbb{R}) \otimes ... \otimes \mathcal{B}(\mathbb{R})}_{d-mal}.$$

Nach Lemma 14.1.13 bilden die offenen Intervalle $\mathcal{I}^\circ_\mathbb{Q} := \{]a, b[ : a, b \in \mathbb{Q}, a \leq b\}$ mit rationalen Endpunkten ein Erzeugendensystem von $\mathcal{B}(\mathbb{R})$ und daher ist nach Lemma 16.1.2

$$\mathcal{B}(\mathbb{R}) \otimes ... \otimes \mathcal{B}(\mathbb{R}) = \sigma(\mathcal{I}^\circ_\mathbb{Q} \times ... \times \mathcal{I}^\circ_\mathbb{Q}) \subseteq \sigma(\mathcal{E}) = \mathcal{B}(\mathbb{R}^d)$$

□

Wir betrachten nun den Fall zwei Faktoren des Produkts. Die Aussagen für endlich viele Faktoren folgen durch Induktion – so sind die Bezeichnungen einfacher.

**Lemma 16.1.4.** *Es seien $(X_1, \mathcal{A}, \mu_1)$ und $(X_2, \mathcal{A}_2, \mu_2)$ zwei $\sigma$-endliche Maßräume und $f : X_1 \times X_2 \to [0, \infty]$ sei $\mathcal{A}_1 \otimes \mathcal{A}_2$-meßbar. Dann gilt:*

*a) Für jedes $x_1 \in X_1$ ist die partielle Abbildung*

$$\begin{aligned} f(x_1, \cdot) : X_2 &\to [0, \infty] \\ x_2 &\mapsto f(x_1, x_2) \end{aligned}$$

*$\mathcal{A}_2$-meßbar.*

*b) Für jedes $x_1 \in X_1$ ist*

$$\begin{aligned} X_1 &\to [0, \infty] \\ x_1 &\mapsto \int\limits_{X_2} f(x_1, x_2) d\mu_2(x_2) \end{aligned}$$

*$\mathcal{A}_1$-meßbar.*

**Satz 16.1.5.** *Es seien $(X_1, \mathcal{A}_1, \mu_1)$ und $(X_2, \mathcal{A}_2, \mu_2)$ $\sigma$-endliche Maßräume, dann gibt es genau ein Produktmaß $\mu = \mu_1 \otimes \mu_2$ auf $\mathcal{A}_1 \otimes \mathcal{A}_2$, dh. genau ein Maß auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ mit $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$ für alle $A = A_1 \times A_2 \in \mathcal{Z}$. Für $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ ist*

$$(\mu_1 \otimes \mu_2)(A) = \int\limits_{X_1} \left( \int\limits_{X_2} \mathbf{1}_A(x_1, x_2) d\mu_2(x_2) \right) d\mu_1(x_1) = \int\limits_{X_2} \left( \int\limits_{X_1} \mathbf{1}_A(x_1, x_2) d\mu_1(x_1) \right) d\mu_2(x_2) \tag{16.1.4}$$

*Beweis.* Es sei $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$, dann ist nach Lemma 16.1.4 die Funktion

$$\begin{array}{rcl} X_1 & \to & [0,\infty] \\ x_1 & \mapsto & \int\limits_{X_2} \mathbf{1}_A(x_1,\cdot)d\mu_2 \end{array}$$

$\mathcal{A}_1$-meßbar, also ist

$$\mu_1 \otimes \mu_2(A) := \int\limits_{X_1} \left( \int\limits_{X_2} \mathbf{1}_A(x_1,x_2)d\mu_2(x_2) \right) d\mu_1(x_1) \in [0,\infty]$$

als Integral einer nichtnegativen meßbaren Funktion in $[0,\infty]$ wohldefiniert. Sind $A, B \in \mathcal{A}_1 \otimes \mathcal{A}_2$ paarweise disjunkt, also $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$, so folgt aus der Linearität des Integrals

$$\mu_1 \otimes \mu_2(A \cup B) = \mu_1 \otimes \mu_2(A) + \mu_1 \otimes \mu_2(B), \tag{16.1.5}$$

dh. $\mu_1 \otimes \mu_2$ ist additiv. Für eine Folge $(B_n)_{n\in\mathbb{N}}$ in $\mathcal{A}_1 \otimes \mathcal{A}_2$ mit $B_n \subseteq B_{n+1}$ und $B := \bigcup\limits_{n\in\mathbb{N}} B_n$ gilt $\mathbf{1}_{B_n} \nearrow \mathbf{1}_B$ und daher folgt aus dem Satz über monotone Konvergenz

$$\mu_1 \otimes \mu_2(B_n) \nearrow \mu_1 \otimes \mu_2(B)$$

und damit ist $\mu_1 \otimes \mu_2$ $\sigma$-stetig von unten, also nach Lemma 14.2.5 ein Maß.

$$\begin{aligned} \mu_1 \otimes \mu_2(A_1 \times A_2) &= \int\limits_{X_1} \left( \int\limits_{X_2} \mathbf{1}_{A_1\times A_2}(x_1,x_2)d\mu_2(x_2) \right) d\mu_1(x_1) \\ &= \int\limits_{X_1} \mu_2(A_2)\mathbf{1}_{A_1}(x_1)d\mu_1(x_1) = \mu_1(A_1)\mu_2(A_2) \end{aligned}$$

ergibt die behaupteten Werte auf Zylindermengen $A_1 \times A_2 \in \mathcal{Z}$. Insbesondere ist $\mu_1 \otimes \mu_2$ ein $\sigma$-endliches Maß, wie es die Wahl von Folgen $(Y_n)_{n\in\mathbb{N}}$ in $\mathcal{A}_1$ mit $\mu_1(Y_n) < \infty$ und $\bigcup\limits_{n\in\mathbb{N}} Y_n = X_1$ bzw. von $(Z_n)_{n\in\mathbb{N}}$ in $\mathcal{A}_2$ mit $\mu_2(Z_n) < \infty$ und $\bigcup\limits_{n\in\mathbb{N}} Z_n = X_2$ wegen $\bigcup\limits_{n\in\mathbb{N}} Y_n \times Z_n = X_1 \times X_2$ und $\mu_1 \otimes \mu_2(Y_n \times Z_n) = \mu_1(Y_n)\mu_2(Z_n) < \infty$ zeigt. Nach dem Eindeutigkeitssatz 14.2.8 für $\sigma$-endliche Maße ist $\mu_1 \otimes \mu_2$ durch die Bedingung $\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$, $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$ auf $\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{Z})$ eindeutig bestimmt. Analog zeigt man durch Vertauschen von $X_1$ und $X_2$, daß auch durch

$$\tilde{\mu}(A) := \int\limits_{X_2} \left( \int\limits_{X_1} \mathbf{1}_A(x_1,x_2)d\mu_1(x_1) \right) d\mu_2(x_2)$$

ein $\sigma$-endliches Maß auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ definiert wird mit $\tilde{\mu}(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$ für alle $A_1 \in \mathcal{A}_1$, womit der Eindeutigkeitssatz für $\sigma$-endliche Maße $\mu_1 \otimes \mu_2 = \tilde{\mu}$ zeigt. □

**Definition 16.1.6.** $\lambda^d := \underbrace{\lambda \otimes ... \otimes \lambda}_{d-mal} : \mathcal{B}(\mathbb{R}^d) \to [0,\infty]$ *heißt das Borel-Lebesguesche Maß auf* $\mathbb{R}^d$. *Die Vervollständigung* $\widetilde{\lambda^d} : \widetilde{\mathcal{B}(\mathbb{R}^d)} \to [0\infty]$ *heißt das Lebesguemaß auf* $\mathbb{R}^d$.

**Korollar 16.1.7** (Cavalieri-Prinzip). *Zu $d \in \mathbb{N}$, $d \geq 2$, $A \in \mathcal{B}(\mathbb{R}^d)$ und $t \in \mathbb{R}$ ist die Schnittmenge*

$$A_{[t]} := \{(x_2, ..., x_d) \in \mathbb{R}^{d-1} : (t, x_2, ..., x_d) \in A\} \in \mathcal{B}(\mathbb{R}^{d-1})$$

*und*

$$\lambda^d(A) = \int\limits_{\mathbb{R}} \lambda^{d-1}(A_{[t]}) d\lambda(t) \tag{16.1.6}$$

*Beweis.* Nach Lemma 16.1.4 ist mit $\mathbf{1}_A$ auch die partielle Abbildung

$$\begin{array}{rcl} \mathbf{1}_A(t, \cdot) : \mathbb{R}^{d-1} & \to & [0, \infty] \\ (x_2, ..., x_d) & \mapsto & \mathbf{1}_A(t, x_2, ..., x_d) \end{array}$$

meßbar und damit $A_{[t]} = (\mathbf{1}_A(t, \cdot))^{-1}(\{1\}) \in \mathcal{B}(\mathbb{R}^{d-1})$. Satz 16.1.5 angewandt auf $\mu_1 = \lambda^{d-1}$ und $\mu_2 = \lambda$ zeigt

$$\lambda^d(A) = \int\limits_{\mathbb{R}} \lambda^{d-1}(A_{[t]}) d\lambda(t).$$

□

**Bemerkung 16.1.8.** Mit dem Cavalieri-Prinzip lassen sich für einfache Gebilde wie Kreise, Kugeln, Zylinder, Rotationskörper induktiv Flächeninhalte bzw. Volumina berechnen. Zum Beispiel für einen Kreis $K := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ um $(0, 0)$ mit Radius 1 ist für $x \in \mathbb{R}$ die Schnittmenge $K_{[x]} = \begin{cases} [-\sqrt{1-x^2}, \sqrt{1-x^2}] & \text{für } x \in [-1, 1] \\ \emptyset & \text{für } x \in \mathbb{R} \backslash [-1, 1] \end{cases}$, also $\lambda(K_{[x]}) = \begin{cases} 2\sqrt{1-x^2} & \text{für } x \in [-1, 1] \\ 0 & \text{für } x \in \mathbb{R} \backslash [-1, 1] \end{cases}$ und damit $\lambda^2(K) = 2 \int\limits_{-1}^{1} \sqrt{1-x^2} dx$.

**Satz 16.1.9** (Tonelli). *Es seien $(X_1, \mathcal{A}_1, \mu_1)$ und $(X_2, \mathcal{A}_2, \mu_2)$ $\sigma$-endliche Maßräume und $f : X_1 \times X_2 \to [0, \infty]$ sei $\mathcal{A}_1 \otimes \mathcal{A}_2$-meßbar. Dann ist*

$$\begin{array}{rcll} X_1 & \to & [0, \infty] & \mathcal{A}_1 - \textit{meßbar} \\ x_1 & \mapsto & \int\limits_{X_2} f(x_1, x_2) d\mu_2(x_2) & \end{array} \tag{16.1.7}$$

$$\begin{array}{rcll} X_2 & \to & [0, \infty] & \mathcal{A}_2 - \textit{meßbar} \\ x_2 & \mapsto & \int\limits_{X_1} f(x_1, x_2) d\mu_1(x_1) & \end{array} \tag{16.1.8}$$

*und es gilt:*

$$\begin{aligned} \int\limits_{X_1 \times X_2} f(x_1, x_2) d(\mu_1 \otimes \mu_2)(x_1, x_2) &= \int\limits_{X_1} \left( \int\limits_{X_2} f(x_1, x_2) d\mu_2(x_2) \right) d\mu_1(x_1) \\ &= \int\limits_{X_2} \left( \int\limits_{X_1} f(x_1, x_2) d\mu_1(x_1) \right) d\mu_2(x_2) \end{aligned} \tag{16.1.9}$$

*Beweis.* Nach Lemma 16.1.4 folgt (16.1.7) und nach Vertauschen der Rollen von $X_1$ und $X_2$ ergibt Lemma 16.1.4 genauso (16.1.8). Die iterierten Integrale

$$\int_{X_1}\left(\int_{X_2} f(x_1,x_2)d\mu_2(x_2)\right)d\mu_1(x_1) \quad \text{und} \quad \int_{X_2}\left(\int_{X_1} f(x_1,x_2)d\mu_1(x_1)\right)d\mu_2(x_2)$$

sind daher als Integrale nichtnegativer, meßbarer Funktionen in $[0,\infty]$ wohldefiniert. Nach Satz 16.1.5 gilt (16.1.9) für $f = \mathbf{1}_A$ mit $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$, woraus wegen Linearität des Integrals (16.1.9) für alle nichtnegativen $\mathcal{A}_1 \otimes \mathcal{A}_2$-Stufenfunktionen $f$ folgt. Ist nun $f : X_1 \times X_2 \to [0,\infty]$ $\mathcal{A}_1 \otimes \mathcal{A}_2$-meßbar, so wähle eine Folge $(f_n)_{n\in\mathbb{N}}$ von nichtnegativen $\mathcal{A}_1 \otimes \mathcal{A}_2$-Stufenfunktionen mit $f_n \nearrow f$ und aus dem Satz über monotone Konvergenz folgt:

$$\begin{aligned}
&\int_{X_1\times X_2} f_n(x_1,x_2)d(\mu_1\otimes\mu_2)(x_1,x_2) \nearrow \int_{X_1\times X_2} f(x_1,x_2)d(\mu_1\otimes\mu_2)(x_1,x_2)\\
=&\int_{X_1}\left(\int_{X_2} f_n(x_1,x_2)d\mu_2(x_2)\right)d\mu_1(x_1) \nearrow \int_{X_1}\left(\int_{X_2} f(x_1,x_2)d\mu_2(x_2)\right)d\mu_1(x_1)\\
=&\int_{X_2}\left(\int_{X_1} f_n(x_1,x_2)d\mu_1(x_1)\right)d\mu_2(x_2) \nearrow \int_{X_2}\left(\int_{X_1} f(x_1,x_2)d\mu_1(x_1)\right)d\mu_2(x_2)
\end{aligned}$$

was (16.1.9) zeigt. □

**Satz 16.1.10** (Fubini)**.** *Es seien $(X_1,\mathcal{A}_1,\mu_1)$ und $(X_2,\mathcal{A}_2,\mu_2)$ $\sigma$-endliche Maßräume, $f : X_1\times X_2 \to \hat{\mathbb{C}}$ sei $\mu_1\otimes\mu_2$-integrierbar. Dann sind die partiellen Abbildungen*

$$\begin{array}{rcl} f(x_1,\cdot) : X_2 & \to & \hat{\mathbb{C}} \\ x_2 & \mapsto & f(x_1,x_2) \end{array}$$

*für $\mu_1$-fast alle $x_1 \in X_1$ $\mu_2$-integrierbar und*

$$\begin{array}{rcl} f(\cdot,x_2) : X_1 & \to & \hat{\mathbb{C}} \\ x_1 & \mapsto & f(x_1,x_2) \end{array}$$

*für $\mu_2$-fast alle $x_2 \in X_2$ $\mu_1$-integrierbar und es gilt:*

$$\int_{X_1\times X_2} fd(\mu_1\otimes\mu_2) = \int_{X_1}\left(\int_{X_2} f(x_1,\cdot)d\mu_2\right)d\mu_1(x_1) = \int_{X_2}\left(\int_{X_1} f(\cdot,x_2)d\mu_1\right)d\mu_2(x_2). \tag{16.1.10}$$

*Beweis.* Nach dem Satz von Tonelli ist

$$\int_{X_1}\left(\int_{X_2} |f(x_1,\cdot)|d\mu_2\right)d\mu_1(x_1) = \int_{X_2}\left(\int_{X_1} |f(\cdot,x_2)|d\mu_1\right)d\mu_2(x_2) = \int_{X_1\times X_2} |f|d(\mu_1\otimes\mu_2) < \infty$$

Lemma 14.3.11 zeigt daher, daß

$$\mu_1\left(\left\{x_1 \in X_1 : \int\limits_{X_2} |f(x_1,\cdot)| d\mu_2 = \infty\right\}\right) = \mu_2\left(\left\{x_2 \in X_2 : \int\limits_{X_1} |f(\cdot,x_2)| d\mu_1 = \infty\right\}\right) = 0$$

Es sei nun $N_1 \in \mathcal{A}_1$ mit $\mu_1(N_1) = 0$, so daß $f(x_1,\cdot)$ für alle $x_1 \in X_1 \backslash N_1$ $\mu_2$-integrierbar ist. Dann sind wegen $|\mathrm{Re}\, f_\pm(x_1,\cdot)|, |\mathrm{Im}\, f_\pm(x_1,\cdot)| \leq |f(x_1,\cdot)|$ auch $\mathrm{Re}\, f_\pm(x_1,\cdot)$ und $\mathrm{Im}\, f_\pm(x_1,\cdot)$ für alle $x_1 \in X_1 \backslash N_1$ $\mu_2$-integrierbar. Damit ist nach dem Satz von Tonelli

$$\begin{aligned} \int\limits_{X_1 \times X_2} \mathrm{Re}\, f_+ d(\mu_1 \otimes \mu_2) &= \int\limits_{(X_1\backslash N_1)\times X_2} \mathrm{Re}\, f_+ d(\mu_1 \otimes \mu_2) = \int\limits_{X_1 \backslash N_1} \left(\int\limits_{X_2} \mathrm{Re}\, f_+(x_1,\cdot) d\mu_2\right) d\mu_1(x_1) \\ &= \int\limits_{X_1} \left(\int\limits_{X_2} \mathrm{Re}\, f_+(x_1,\cdot) d\mu_2\right) d\mu_1(x_1) < \infty \qquad (16.1.11) \end{aligned}$$

und die analogen Gleichungen gelten auch für $\mathrm{Re}\, f_-$ und $\mathrm{Im}\, f_\pm$. Nach Zerlegung in Real- und Imaginärteil und dann weiter in Postiv- und Negativteil folgt unter Verwendung von (16.1.11) bzw. den analogen Gleichungen für $\mathrm{Re}\, f_-$ bzw. $\mathrm{Im}\, f_\pm$:

$$\begin{aligned} \int\limits_{X_1 \times X_2} f d(\mu_1 \otimes \mu_2) &= \int\limits_{(X_1\backslash N_1)\times X_2} f d(\mu_1 \otimes \mu_2) = \int\limits_{X_1 \backslash N_1} \left(\int\limits_{X_2} f(x_1,\cdot) d\mu_2\right) d\mu_1(x_1) \\ &= \int\limits_{X_1} \left(\int\limits_{X_2} f(x_1,\cdot) d\mu_2\right) d\mu_1(x_1) \end{aligned}$$

Die andere Gleichung folgt analog durch Vertauschen von $X_1$ und $X_2$. □

**Bemerkung 16.1.11.**

a) Im Satz von Fubini ist die Voraussetzung „$f$ ist $\mu_1 \otimes \mu_2$-integrierbar“ wesentlich, wie das folgende Beispiel zeigt: Betrachte die Mengen

$$\begin{aligned} A_- &:= \{(x,y) \in \mathbb{R}^2 : y > 0, x < y < x+1\} \\ A_+ &:= \{(x,y) \in \mathbb{R}^2 : Y > 0, x-1 < y < x\} \end{aligned}$$

und $f := \mathbf{1}_{A_+} - \mathbf{1}_{A_-}$. Dann ist $\lambda^2(A_+) = \lambda^2(A_-) = \infty$ und damit $f$ nicht $\lambda^2-$integrierbar. Für jedes $y > 0$ ist

$$\int\limits_{\mathbb{R}} f(x,y) d\lambda(x) = \lambda(]y,y+1[) - \lambda(]y-1,y[) = 0,$$

also

$$\int\limits_{\mathbb{R}} \left(\int\limits_{\mathbb{R}} f(x,y) d\lambda(x)\right) d\lambda(y) = 0.$$

Ebenso ist $\int\limits_{\mathbb{R}} f(x,y)d\lambda(y) = 0$ für $x > 1$ und für $x \in [-1,1]$ ist

$$\int\limits_{\mathbb{R}} f(x,y)d\lambda(y) = \max\{x,0\} - \min\{x+1,1\}$$

und damit

$$\int\limits_{\mathbb{R}} \left( \int\limits_{\mathbb{R}} f(x,y)d\lambda(y) \right) d\lambda(x) = \int\limits_{[-1,1]} \left( \int\limits_{\mathbb{R}} f(x,y)d\lambda(y) \right) d\lambda(x) = -1$$

b) Wenn nicht $(X_1, \mathcal{A}_1, \mu_1)$ und $(X_2, \mathcal{A}_2, \mu_2)$ beide $\sigma$-endlich sind, dann werden die Sätze von Tonelli und Fubini falsch: Für $(X_1, \mathcal{A}_1, \mu_1) = ([0,1], \mathcal{B}([0,1]), \lambda)$ und $(X_2, \mathcal{A}_2, \mu_2) = ([0,1], \mathcal{B}([0,1]), \nu)$ ist das Zählmaß $\nu$ auf $\mathcal{B}([0,1])$ nicht $\sigma$-endlich und für $f := \mathbf{1}_{\Gamma(\mathrm{id}_{[0,1]})} : \begin{array}{rcl} [0,1]^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & \begin{cases} 1 & \text{für } x = y \\ 0 & \text{für } x \neq y \end{cases} \end{array}$ ist als Indikatorfunktion der abgeschlossenen Menge $\{(x,x) \in [0,1]^2 : x \in [0,1]\}$ wieder meßbar, aber

$$\int\limits_{[0,1]} \left( \int\limits_{[0,1]} f(x,y)d\nu(y) \right) d\lambda(x) = \int\limits_{[0,1]} 1\, d\lambda = 1$$

$$\int\limits_{[0,1]} \left( \int\limits_{[0,1]} f(x,y)d\lambda(x) \right) d\nu(y) = \int\limits_{[0,1]} 0\, d\nu = 0$$

## 16.2 Der Transformationssatz

**Lemma 16.2.1.** *$\lambda^d$ ist ein translationsinvariantes Maß, dh. für* $T_{\underline{a}} : \begin{array}{rcl} \mathbb{R}^d & \to & \mathbb{R}^d \\ \underline{x} & \mapsto & \underline{x} + \underline{a} \end{array}$ *und jedes $\underline{a} \in \mathbb{R}^d$ gilt:*

$$T_{\underline{a}}(\mathcal{B}(\mathbb{R}^d)) = \mathcal{B}(\mathbb{R}^d) \quad \text{und} \quad \lambda^d \circ T_{\underline{a}} = \lambda^d. \tag{16.2.1}$$

*Ist $\mu : \mathcal{B}(\mathbb{R}^d) \to [0,\infty]$ ein translationsinvariantes Maß mit $\mu(]0,1]^d) := c_\mu < \infty$, dann gilt $\mu = c_\mu \lambda^d$*

*Beweis.* Für jedes $\underline{a} \in \mathbb{R}^d$ ist $T_{\underline{a}} \circ T_{-\underline{a}} = \mathrm{id}_{\mathbb{R}^d}$, und $T_{-\underline{a}} : \mathbb{R}^d \to \mathbb{R}^d$ stetig, also nach Korollar 14.1.12 Borel-meßbar und damit $(T_{-\underline{a}})^{-1}(B) = T_{\underline{a}}(B) \in \mathcal{B}(\mathbb{R}^d)$ für jedes $\underline{a} \in \mathbb{R}^d$ und $B \in \mathcal{B}(\mathbb{R}^d)$. Dies zeigt $T_{\underline{a}}(\mathcal{B}(\mathbb{R}^d)) \subseteq \mathcal{B}(\mathbb{R}^d)$ und wegen $T_{\underline{a}} \circ T_{-\underline{a}} = \mathrm{id}_{\mathbb{R}^d}$ ist

$$\mathcal{B}(\mathbb{R}^d) = (T_{-\underline{a}} \circ T_{\underline{a}})(\mathcal{B}(\mathbb{R}^d)) = T_{-\underline{a}}(T_{\underline{a}}(\mathcal{B}(\mathbb{R}^d)) \subseteq T_{\underline{a}}(\mathcal{B}(\mathbb{R}^d)) \subseteq \mathcal{B}(\mathbb{R}^d),$$

was insbesondere $T_{\underline{a}}(\mathcal{B}(\mathbb{R}^d)) = \mathcal{B}(\mathbb{R}^d)$ zeigt. $\lambda^d \circ T_{\underline{a}} = \lambda^d \circ (T_{-\underline{a}})^{-1}$ ist ein Bildmaß mit

$$\begin{aligned}(\lambda^d \circ T_{\underline{a}})(]x_1, y_1[\times ... \times ]x_d, y_d[) &= \lambda^d(]x_1 + a_1, y_1 + a_1[\times ... \times ]x_d + a_d, y_d + a_d[) \\ &= \prod_{k=1}^{d} (y_k - x_k) = \lambda^d(]x_1, y_1[\times ... \times ]x_d, y_d[)\end{aligned}$$

für jedes $\underline{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_d \end{pmatrix} \in \mathbb{R}^d$. $\lambda^d$ und $\lambda^d \circ T_{\underline{a}}$ sind $\sigma$-endliche Maße, denn zB. ist $\mathbb{R}^d = \bigcup_{n \in \mathbb{N}} ]-n, n[^d$ und $\lambda^d \circ T_{\underline{a}}(]-n, n[^d) = \lambda^d(]-n, n[^d) = (2n)^d$. Auf dem Erzeugendensystem

$$\mathcal{E} := \{]x_1, y_1[\times ... \times ]x_d, y_d[ : x_1, ..., x_d, y_1, ..., y_d \in \mathbb{R}, x_1 \leq y_1, ..., x_d \leq y_d\}$$

gilt $\lambda^d \circ T_{\underline{a}}|_{\mathcal{E}} = \lambda^d|_{\mathcal{E}}$ und damit folgt $\lambda^d \circ T_{\underline{a}} = \lambda^d$ aus dem Eindeutigkeitssatz 14.2.8. Zu $(n_1, ..., n_d) \in \mathbb{N}^d$ betrachte das Gitter

$$G_{(n_1,...,n_d)} := \left\{ \left( \frac{k_1}{n_1}, ..., \frac{k_d}{n_d} \right) \in \mathbb{Q}^d : k_1, ..., k_d \in \mathbb{N}, 0 \leq k_j < n_j \text{ für alle } j = 1, ..., d \right\}.$$

Verschieben des halboffenen Rechtecks $]0, \frac{1}{n_1}] \times ... \times ]0, \frac{1}{n_d}]$ um jeden Punkt von $G_{(n_1,...,n_d)}$ ergibt

$$]0, 1]^d = \bigcup_{(\frac{k_1}{n_1}, ..., \frac{k_d}{n_d}) \in G_{(n_1,...,n_d)}} \left] \frac{k_1}{n_1}, \frac{k_1 + 1}{n_1} \right] \times ... \times \left] \frac{k_d}{n_d}, \frac{k_d + 1}{n_d} \right]$$

als Vereinigung von $n_1 \cdots n_d$ paarweise disjunkten Mengen, von denen jede wegen Translationsinvarianz von $\mu$ alle gleiches Maß besitzen. Wegen $\mu(]0, 1]^d) = c_\mu$ folgt

$$\mu \left( \left] \frac{k_1}{n_1}, \frac{k_1 + 1}{n_1} \right] \times ... \times \left] \frac{k_d}{n_d}, \frac{k_d + 1}{n_d} \right] \right) = c_\mu \frac{1}{n_1} \cdots \frac{1}{n_d}.$$

Die Anwendung der Translationsinvarianz und die Zerlegung $]0, \frac{l_j}{n_j}] = ]0, \frac{1}{n_j}] \cup ... \cup ]\frac{l_1 - 1}{n_j}, \frac{l_j}{n_j}]$ in paarweise disjunkte Mengen ergibt

$$\mu(]0, \frac{l_1}{n_1}] \times ... \times ]0, \frac{l_d}{n_d}]) = c_\mu \frac{l_1}{n_1} \cdots \frac{l_d}{n_d} = c_\mu \lambda^d(]0, \frac{l_1}{n_1}] \times ... \times ]0, \frac{l_d}{n_d}])$$

für jedes $(l_1, ..., l_d) \in \mathbb{N}^d$. Durch Betrachten aller $(n_1, ..., n_d) \in \mathbb{N}^d$ und wegen Translationsinvarianz von $\mu$ und $\lambda^d$ folgt

$$\mu|_{\mathcal{I}_\mathbb{Q} \times ... \times \mathcal{I}_\mathbb{Q}} = c_\mu \lambda^d|_{\mathcal{I}_\mathbb{Q} \times ... \times \mathcal{I}_\mathbb{Q}} \tag{16.2.2}$$

Da die halboffenen Intervalle $\mathcal{I}_\mathbb{Q} := \{]a, b] : a, b \in \mathbb{Q}, a \leq b\}$ mit rationalen Endpunkten nach Lemma 14.1.13 ein durchschnittstabiles Erzeugendensystem von $\mathcal{B}(\mathbb{R})$ bilden, ist

$$\sigma(\mathcal{I}_\mathbb{Q} \times ... \times \mathcal{I}_\mathbb{Q}) = \mathcal{B}(\mathbb{R}^d) = \mathcal{B}(\mathbb{R}) \otimes ... \otimes \mathcal{B}(\mathbb{R})$$

Aufgrund der Translationsinvarianz und der Normierungsbedingung $\mu(]0,1]^d) = c_\mu$ gilt $\mu(]-n,n]^d) = c_\mu \lambda^d(]-n,n]^d) = c_\mu (2n)^d$, also sind wegen $\mathbb{R}^d = \bigcup_{n\in\mathbb{N}} ]-n,n]^d$ die beiden Maße $\mu$ und $c_\mu\lambda^d$ $\sigma$-endlich. Der Eindeutigkeitssatz 14.2.8 für $\sigma$-endliche Maße und (16.2.2) zeigen $c_\mu \lambda^d = \mu$. □

**Lemma 16.2.2.** *Es sei $d \in \mathbb{N}$, $d \geq 2$, $D \in \mathcal{B}(\mathbb{R}^{d-1})$ und $f : D \to \mathbb{R}$ meßbar, dann ist $\lambda^d(\Gamma(f)) = 0$ für den Graphen $\Gamma(f) := \{(x, f(x)) \in \mathbb{R}^d : x \in D\}$ von $f$.*

*Beweis.* Für den Beweis dürfen wir oE. $D = \mathbb{R}^{d-1}$ voraussetzen, denn sonst betrachten wir $\begin{aligned} \tilde{f} : \mathbb{R}^{d-1} &\to \mathbb{R} \\ x &\mapsto \begin{cases} f(x) & \text{für } x \in D \\ 0 & \text{für } x \in \mathbb{R}^{d-1}\setminus D \end{cases}\end{aligned}$ und dann ist $\Gamma(f) = (D \times \mathbb{R}) \cap \Gamma(\tilde{f})$. Die Projektionen $\begin{aligned} P : \mathbb{R}^d &\to \mathbb{R}^{d-1} \\ (x_1, ..., x_d) &\mapsto (x_1, ..., x_{d-1})\end{aligned}$ und $\begin{aligned} Q : \mathbb{R}^d &\to \mathbb{R} \\ (x_1, ..., x_d) &\mapsto x_d\end{aligned}$ sind linear, also nach Satz 13.9.4 stetig und damit nach Korollar 14.1.12 auch meßbar. Damit ist auch

$$\begin{aligned} g := f \circ P - Q : \mathbb{R}^d &\to \mathbb{R} \\ (x_1, ..., x_d) &\mapsto f(x_1, ..., x_{d-1}) - x_d \end{aligned}$$

wieder meßbar, also $\Gamma(f) = \{(x_1, ..., x_d) \in \mathbb{R}^d : x_d = f(x_1, ..., x_{d-1})\} = g^{-1}(\{0\}) \in \mathcal{B}(\mathbb{R}^d)$. Da für alle $(x_1, ..., x_{d-1}) \in \mathbb{R}^{d-1}$ nach Definition einer Funktion

$$\int_{\mathbb{R}} \mathbf{1}_{\Gamma(f)}(x_1, ..., x_{d-1}, t) d\lambda(t) = \lambda(\{f(x_1, ..., x_{d-1})\}) = 0$$

ist, folgt aus dem Satz von Tonelli:

$$\lambda^d(\Gamma(f)) = \int_{\mathbb{R}^{d-1}} \left( \int_{\mathbb{R}} \mathbf{1}_{\Gamma(f)}(x_1, ..., x_{d-1}, t) d\lambda(t) \right) d\lambda^{d-1}(x_1, ..., x_{d-1}) = 0.$$

□

**Definition 16.2.3.** *Es seien $X$ und $Y$ Banachräume über $\mathbb{K}$, $U \subseteq X$ offen, $V \subseteq Y$ offen. Eine Abbildung $f : U \to V$ heißt $\mathcal{C}^1$***-Diffeomorphismus***, wenn*

*a) $f$ bijektiv*

*b) $f$ ist stetig differenzierbar*

*c) Die inverse Abbildung $f^{-1} : V \to U$ ist stetig differenzierbar.*

**Satz 16.2.4** (Transformationssatz)**.** *Es sei $d \in \mathbb{N}$, $U, V \subseteq \mathbb{R}^d$ offen, $X$ ein Banachraum $\psi : U \to V$ ein $\mathcal{C}^1-$Diffeomorphismus und $f : V \to X$ sei meßbar. Dann gilt:*

*a)* $\displaystyle\int_V \|f(v)\| d\lambda^d(v) = \int_U \|(f \circ \psi)(u)\| \cdot |\det D\psi(u)| d\lambda^d(u)$

b) *$f$ ist genau dann $\lambda^d-$integrierbar auf $V$, wenn $(f \circ \psi)$ integrierbar auf $U$ bezüglich $|\det D\psi|\lambda^d$ ist und in diesem Fall gilt:*

$$\int_V f(v)d\lambda^d(v) = \int_U (f \circ \psi)(u)|\det D\psi(u)|d\lambda^d(u). \tag{16.2.3}$$

**Bemerkung 16.2.5.** Im Fall $f = \mathbf{1}_A$ mit $A \in \mathcal{B}(\mathbb{R}^d)$ und für lineares $\psi : \mathbb{R}^d \to \mathbb{R}^d$ gilt

$$\lambda^d(\psi[A]) = |\det(\psi)|\lambda^d(A). \tag{16.2.4}$$

Dies ist für bijektives $\psi$ ein Spezialfall des Transformationssatzes (den man üblicherweise zuerst beweist). Ist $\psi$ nicht bijektiv, also $\det(\psi) = 0$, dann ist $\psi[\mathbb{R}^d]$ ein Unterraum mit $\dim(\psi[\mathbb{R}^d]) \leq d - 1$, also $0 \leq \lambda^d(\psi[A]) \leq \lambda^d(\psi[\mathbb{R}^d]) = 0$ nach Lemma 16.2.2.

Viele Koordinatentransformationen sind keine $\mathcal{C}^1-$Diffeomorphismen, sondern erst nach geeigneter Einschränkung von Definitionsbereich und Wertevorrat ergibt sich ein solcher $\mathcal{C}^1-$Diffeomorphismus. Hat man dabei nur Nullmengen „ignoriert“ dann kann man die folgende Version des Transformationssatzes verwenden.

**Korollar 16.2.6.** *Es sei $d \in \mathbb{N}$, $U \subseteq \mathbb{R}^d$ offen, $\Psi : U \to \mathbb{R}^d$ stetig differenzierbar, $Y \subseteq U$ sei $\lambda^d-$meßbar, $Y \setminus \overset{\circ}{Y}$ sei eine $\lambda^d-$Nullmenge und* $\psi : \begin{array}{ccc} \overset{\circ}{Y} & \to & \Psi(\overset{\circ}{Y}) \\ x & \mapsto & \Psi(x) \end{array}$ *ein $\mathcal{C}^1-$Diffeomorphismus. Für jedes meßbare $f : \Psi(Y) \to X$ gilt:*

a) $$\int_{\Psi(Y)} \|f(v)\|d\lambda^d(v) = \int_Y \|(f \circ \Psi)(u)\| \cdot |\det D\Psi(u)|d\lambda^d(u)$$

b) *$f$ ist genau dann $\lambda^d-$integrierbar, wenn $(f \circ \psi)$ integrierbar bezüglich $|\det D\psi|\lambda^d$ ist und in diesem Fall gilt:*

$$\int_{\Psi(Y)} f(v)d\lambda^d(v) = \int_Y (f \circ \Psi)(u)|\det D\Psi(u)|d\lambda^d(u). \tag{16.2.5}$$

**Lemma 16.2.7.** *Es sei $X$ ein Banachraum, $f : \mathbb{R}^2 \to X$ sei meßbar, dann gilt:*

a) $$\begin{aligned} \int_{\mathbb{R}^2} \|f(x,y)\|d\lambda^2(x,y) &= \int_0^{2\pi}\int_0^{\infty} \|f(r\cos(\varphi), r\sin(\varphi))\|rdrd\varphi \\ &= \int_0^{\infty} r \int_0^{2\pi} \|f(r\cos(\varphi), r\sin(\varphi))\|d\varphi dr \end{aligned}$$

*b) $f$ ist genau dann $\lambda^2-$integrierbar, wenn* $\begin{array}{rcl} g : ]0,\infty[\times]0,2\pi[ & \to & X \\ (r,\varphi) & \mapsto & f(r\cos(\varphi), r\sin(\varphi)) \end{array}$ *integrierbar bezüglich $rdrd\varphi$ ist. In diesem Fall gilt:*

$$\begin{aligned} \int\limits_{\mathbb{R}^2} f(x,y)d\lambda^2(x,y) &= \int\limits_0^{2\pi}\int\limits_0^{\infty} f(r\cos(\varphi), r\sin(\varphi))rdrd\varphi \\ &= \int\limits_0^{\infty} r\int\limits_0^{2\pi} f(r\cos(\varphi), r\sin(\varphi))d\varphi dr \end{aligned}$$

*Beweis.* $Y := ]0,\infty[\times]0,2\pi[$ ist offen und $\begin{array}{rcl} \Psi : \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (r,\varphi) & \mapsto & (r\cos(\varphi), r\sin(\varphi)) \end{array}$ ist stetig differenzierbar. Für $\Psi(Y) = \mathbb{R}^2\backslash([0,\infty[\times\{0\})$ ist $\lambda^2(\mathbb{R}^2\backslash\Psi(Y)) = \lambda^2([0,\infty[\times\{0\}) = 0$. Ferner ist $\begin{array}{rcl} \psi : Y & \mapsto & \Psi(Y) \\ (r,\varphi) & \mapsto & \Psi(r,\varphi) \end{array}$ bijektiv (Polarkoordinaten, vgl. Satz 13.8.8) stetig und die Umkehrfunktion $\begin{array}{rcl} \psi^{-1} : \mathbb{R}^2\backslash\{(x,0) : x \geq 0\} & \to & ]0,\infty[\times]0,2\pi[ \\ (x,y) & \mapsto & (\sqrt{x^2+y^2}, \varphi(x,y)) \end{array}$ mit

$$\varphi(x,y) := \begin{cases} \operatorname{arccot}(\frac{x}{y}) & \text{für } y > 0 \\ \pi & \text{für } y = 0 \\ \pi + \operatorname{arccot}(\frac{x}{y}) & \text{für } y < 0 \end{cases}$$

ist stetig.

$$\det((D\psi)(r,\varphi)) = \det\begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix} = r > 0$$

für alle $(r,\varphi) \in Y$. Damit ist

$$\begin{array}{rcl} (D\psi)(r,\varphi) : \mathbb{R}^2 & \to & \mathbb{R}^2 \\ \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} & \mapsto & (D\psi)(r,\varphi)\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \end{array}$$

für alle $(r,\varphi) \in Y$ ein linearer Homöomorphismus, also die Umkehrfunktion $\psi^{-1}$ nach Satz 15.1.16 in jedem Punkt stetig differenzierbar, also $\psi$ ein $\mathcal{C}^1-$Diffeomorphismus. Nach Korollar 16.2.6 und dem Satz 16.1.9 von Tonelli bzw. dem Satz von Fubini 16.1.10 folgt die Behauptung. □

**Beispiel 16.2.8.**

$$\int\limits_{\mathbb{R}} e^{-x^2}dx = \sqrt{\pi}$$

*Beweis.* Wegen $e^{-x^2} \leq \begin{cases} 1 \text{ für } |x| \leq 1 \\ e^{-|x|} \text{ für } |x| > 1 \end{cases}$ ist $\int\limits_{\mathbb{R}} |e^{-x^2}|dx < \infty$, also ist nach dem Satz von Tonelli

$$\int\limits_{\mathbb{R}^2} e^{-(x^2+y^2)}d\lambda^2(x,y) = \left(\int\limits_{\mathbb{R}} e^{-x^2}dx\right)^2 < \infty.$$

Nach Lemma 16.2.7 ist

$$\int\limits_{\mathbb{R}^2} e^{-(x^2+y^2)} d\lambda^2(x,y) = \int\limits_0^\infty \int\limits_0^{2\pi} e^{-r^2} r d\varphi dr = 2\pi \int\limits_0^\infty r e^{-r^2} dr.$$

Wegen $\mathbf{1}_{[0,N]}(r) r e^{-r^2} \nearrow r e^{-r^2}$ folgt nach dem Satz von der monotonen Konvergenz und dem Hauptsatz der Differential- und Integralrechnung:

$$2\pi \int\limits_0^\infty r e^{-r^2} dr = 2\pi \lim_{N\to\infty} \int\limits_0^N r e^{-r^2} dr = \lim_{N\to\infty} \pi e^{-r^2}\Big|_{r=0}^{r=N} = \pi.$$

$\square$

**Lemma 16.2.9.** *Es sei $X$ ein Banachraum, $f : \mathbb{R}^3 \to X$ sei meßbar, dann gilt:*

*a)*

$$\begin{aligned} &\int\limits_{\mathbb{R}^3} \|f(x,y,z)\| d\lambda^3(x,y,z) \\ = &\int\limits_0^\pi \int\limits_0^{2\pi} \int\limits_0^\infty \|f(r\cos(\varphi)\sin(\theta), r\sin(\varphi)\sin(\theta), r\cos(\theta))\| r^2 dr d\varphi \sin(\theta) d\theta \\ = &\int\limits_0^\infty \int\limits_0^\pi \int\limits_0^{2\pi} \|f(r\cos(\varphi)\sin(\theta), r\sin(\varphi)\sin(\theta), r\cos(\theta))\| d\varphi \sin(\theta) d\theta r^2 dr \end{aligned}$$

*b) $f$ ist genau dann $\lambda^3-$integrierbar, wenn*

$$\begin{array}{rcl} g : ]0,\infty[\times]0,2\pi[\times]0,\pi[ & \to & X \\ (r,\varphi,\theta) & \mapsto & f(r\cos(\varphi)\sin(\theta), r\sin(\varphi)\sin(\theta), r\cos(\theta)) \end{array}$$

*integrierbar bezüglich $r^2 dr d\varphi \sin(\theta) d\theta$ ist. In diesem Fall gilt:*

$$\begin{aligned} &\int\limits_{\mathbb{R}^3} f(x,y,z) d\lambda^3(x,y,z) \\ = &\int\limits_0^\pi \int\limits_0^{2\pi} \int\limits_0^\infty f(r\cos(\varphi)\sin(\theta), r\sin(\varphi)\sin(\theta), r\cos(\theta)) r^2 dr d\varphi \sin(\theta) d\theta \\ = &\int\limits_0^\infty \int\limits_0^\pi \int\limits_0^{2\pi} f(r\cos(\varphi)\sin(\theta), r\sin(\varphi)\sin(\theta), r\cos(\theta)) d\varphi \sin(\theta) d\theta r^2 dr \end{aligned}$$

## 16.3 Die Banachräume $L^p(X)$

**Bemerkung 16.3.1.** Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Für jede meßbare Funktion $f : X \to \widehat{\mathbb{R}}$ oder $f : X \to \widehat{\mathbb{C}}$ mit $\int_X |f| d\mu < \infty$ ist nach Lemma 14.3.11 die Menge $\{x \in X : |f(x)| = \infty\}$ eine $\mu$-Nullmenge. In

$$\mathcal{L}^1(X, \widehat{\mathbb{C}}) := \{f : X \to \widehat{\mathbb{C}} : f \text{meßbar}, \int_X |f| d\mu < \infty\} \tag{16.3.1}$$

ist

$$\mathcal{N}(X, \widehat{\mathbb{C}}) := \{f \in \mathcal{L}^1(X, \widehat{\mathbb{C}}) : f = 0 \ \mu - \text{fast überall }\} \tag{16.3.2}$$

ein Untervektorraum. Auf $\mathcal{L}^1(X)$ definiert $f \sim g$ genau dann wenn $f - g \in \mathcal{N}(X)$ offenbar eine Äquivalenzrelation.

**Definition 16.3.2.** *Es sei $p \in [1, \infty[$, dann heißt*

$$\mathcal{L}^p(X) := \{f : X \to \widehat{\mathbb{C}} : f \text{meßbar}, \int_X |f|^p d\mu < \infty\} \tag{16.3.3}$$

*die Menge der p-fach integrierbaren Funktionen und*

$$L^p(X) := \mathcal{L}^p(X)/\mathcal{N}(X, \widehat{\mathbb{C}}) \tag{16.3.4}$$

*die Menge aller Restklassen $\mu$-fast überall gleicher p-fach integrierbarer Funktionen. Hat man auf einer Menge verschiedene Maße, so schreibt man $\mathcal{L}^p(\mu)$ statt $\mathcal{L}^p(X)$ zur Klarstellung bezüglich welchen Maßes integriert wird.*

**Bemerkung 16.3.3.** Für $p \in [1, \infty[$ ist

$$\|f\|_{L^p} := \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}} \tag{16.3.5}$$

unabhängig von der Wahl des Repräsentanten $f$ in der Restklasase $[f]_\sim \in L^p(X)$.

*Beweis.* Sind $f_1, f_2 \in [f]_\sim$, dh. $f_1 = f$ $\mu$-fast überall und $f_2 = f$ $\mu$-fast überall, dann ist $f_1 = f_2$ $\mu$-fast überall, also

$$\begin{aligned} \|f_1\|_{L^p}^p - \|f_2\|_{L^p}^p &= \int_X (|f_1|^p(x) - |f_2|^p(x))\, d\mu(x) = \int_{\{x \in X : |f_1(x)| \neq |f_2(x)|\}} \big(|f_1|^p(x) - |f_2|^p(x)\big) d\mu(x) \\ &= -\int_{\{x \in X : |f_1(x)| < |f_2(x)|\}} \big(|f_2|^p(x) - |f_1|^p(x)\big) d\mu(x) + \int_{\{x \in X : |f_1(x)| > |f_2(x)|\}} \big(|f_1|^p(x) - |f_2|^p(x)\big) d\mu(x) \\ &= 0, \end{aligned}$$

nach Lemma 14.3.11, denn $(|f_2|^p - |f_1|^p)\mathbf{1}_{\{|f_1| < |f_2|\}}$ und $(|f_1|^p - |f_2|^p)\mathbf{1}_{\{|f_1| > |f_2|\}}$ sind meßbare $[0, \infty]$-wertige Funktionen und $\{|f_1| < |f_2|\}$, $\{|f_2| < |f_1|\} \subseteq \{|f_1| \neq |f_2|\}$ sind $\mu$-Nullmengen. □

**Definition 16.3.4.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum, dann ist*

$$\mathcal{L}^\infty(X, \widehat{\mathbb{C}}) := \Big\{ f : X \to \widehat{\mathbb{C}} : f \text{ meßbar}, \\ \|f\|_{L^\infty} := \inf \big\{ \sup\{|f(x)| : x \in A\} : A \in \mathcal{A}, \mu(X \backslash A) = 0 \big\} < \infty \Big\}$$

*die Menge aller $\mu$-fast überall beschränkten meßbaren Funktionen, bzw.*

$$L^\infty(X) := \mathcal{L}^\infty(X, \widehat{\mathbb{C}}) / \mathcal{N}(X, \widehat{\mathbb{C}}) \tag{16.3.6}$$

*die Restklassen $\mu$-fast überall gleicher, $\mu$-fast überall beschränkter meßbarer Funktionen.*

**Lemma 16.3.5.** *Es sei $p \in [1, \infty[$, $f \in L^p(X)$ und $g \in L^\infty(X)$, dann ist $fg \in L^p(X)$.*

*Beweis.* Wegen $g \in L^\infty(X)$ ist $\|g\|_{L^\infty} < \infty$ und $|g(x)| \leq \|g\|_{L^\infty}$ für $\mu$-fast alle $x \in X$. Damit ist $|f(x)g(x)| \leq \|g\|_{L^\infty} |f(x)|$ für $\mu$-fast alle $x \in X$, also folgt nach Bilden der $p$-ten Potenz und Integration

$$\int_X |f(x)g(x)|^p d\mu(x) \leq \|g\|^p_{L^\infty} \int_X |f(x)|^p d\mu(x) = \|g\|^p_{L^\infty} \|f\|^p_p$$ □

**Satz 16.3.6** (Hölder-Ungleichung)**.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum, $p, q \in ]1, \infty[$ mit $1 = \frac{1}{p} + \frac{1}{q}$ und $f \in L^p(X)$, $g \in L^q(X)$, dann ist $fg \in L^1(X)$ und*

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \|g\|_{L^q} \tag{16.3.7}$$

*Beweis.* Offenbar ist für jeden Repräsentanten $f \in \mathcal{L}^p(X)$ und $g \in \mathcal{L}^q(X)$ auch das punktweise Produkt $fg : X \to \widehat{\mathbb{C}}$ wieder meßbar. Ist $\|f\|_{L^p} = 0$ oder $\|g\|_{L^q} = 0$, dann ist $|f|^p = 0$ $\mu$-fast überall bzw. $|g|^q = 0$ $\mu$-fast überall nach Lemma 14.3.11, also $fg = 0$ $\mu$-fast überall und damit $\|fg\|_{L^1} = 0$. Wir dürfen also für den Beweis $\|f\|_{L^p} \neq 0$ und $\|g\|_{L^q} \neq 0$ voraussetzen. Für alle $a, b \in ]0, \infty[$ folgt aus der Konkavität des ln – dh. für alle $s, t \in ]0, \infty[$ und alle $r \in [0, 1]$ gilt:

$$\ln(rt + (1 - r)s) \geq (1 - r)\ln(s) + r\ln(t)$$

folgt wegen $\frac{1}{q} = 1 - \frac{1}{p}$:

$$\begin{aligned} \ln\left(\frac{a^p}{p} + \frac{b^q}{q}\right) &= \ln\left(\frac{1}{p}a^p + (1 - \frac{1}{p})b^{\frac{p}{p-1}}\right) \geq \frac{1}{p}\ln(a^p) + (1 - \frac{1}{p})\ln(b^{\frac{p}{p-1}}) \\ &= \ln(a) + \ln(b) = \ln(ab). \end{aligned}$$

Da $\exp : \mathbb{R} \to ]0, \infty[$ streng monoton steigend und die Umkehrfunktion von ln ist, folgt:

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q \tag{16.3.8}$$

für $a, b \in ]0, \infty[$ und (16.3.8) gilt offenbar auch für $a = 0$ oder $b = 0$. Ferner zeigt (16.3.8) auch

$$\{x \in X : |f(x)g(x)| = \infty\} \subseteq \{x \in X : |f(x)|^p = \infty\} \cup \{x \in X : |g(x)|^q = \infty\} =: N.$$

Wegen $\int\limits_X |f|^p d\mu < \infty$ und $\int\limits_X |g|^q d\mu < \infty$ ist $N$ nach Lemma 14.3.11 eine $\mu$-Nullmenge, daher folgt aus der Monotonie des Integrals

$$\begin{aligned}\frac{\|fg\|_{L^1}}{\|f\|_{L^p}\|g\|_{L^q}} &= \frac{1}{\|f\|_{L^p}\|g\|_{L^q}}\int\limits_X |fg|d\mu = \int\limits_X \mathbf{1}_{X\setminus N}\frac{|f|}{\|f\|_{L^p}}\frac{|g|}{\|g\|_{L^q}}d\mu \\ &\le \frac{1}{p\|f\|_{L^p}^p}\int\limits_X \mathbf{1}_{X\setminus N}|f|^p d\mu + \frac{1}{q\|g\|_{L^q}^q}\int\limits_X \mathbf{1}_{X\setminus N}|g|^q d\mu \\ &= \frac{1}{p\|f\|_{L^p}^p}\int\limits_X |f|^p d\mu + \frac{1}{q\|g\|_{L^q}^q}\int\limits_X |g|^q d\mu = \frac{1}{p}+\frac{1}{q} = 1\end{aligned}$$

dh. $\|fg\|_{L^1} \le \|f\|_{L^p}\|g\|_{L^q}$, also $fg \in \mathcal{L}^1(X)$ für die Repräsentanten oder $fg \in L^1(X)$. $\square$

**Lemma 16.3.7.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum, $p \in [1,\infty]$, dann ist $(L^p(X), \|\cdot\|_{L^p})$ ein normierter Raum.*

*Beweis.* Es seien $f, g \in \mathcal{L}^p(X)$ Repräsentanten von $f, g \in L^p(X)$. Dann ist $f + \lambda g : X \to \widehat{\mathbb{C}}$ für jedes $\lambda \in \mathbb{C}$ meßbar. Die Behauptung folgt für $p = \infty$ durch Einsetzen in die Definition. Wir betrachten also nun den Fall $p \in [1,\infty[$. Da $\begin{array}{ccc}[0,\infty[ & \to & [0,\infty[\\ x & \mapsto & x^p\end{array}$ für jedes $p \in [1,\infty[$ monoton steigend ist, folgt

$$(a+b)^p \le (2\max\{a,b\})^p = 2^p \max\{a^p, b^p\} \le 2^p(a^p + b^p) \tag{16.3.9}$$

für alle $a, b \in [0,\infty[$ und (16.3.9) gilt offenbar auch für $a = \infty$ oder $b = \infty$. Daraus folgt

$$\begin{aligned}\|f+\lambda g\|_{L^p}^p &= \int\limits_X |f(x)+\lambda g(x)|^p d\mu(x) \le \int\limits_X 2^p\Big(|f(x)|^p + |\lambda|^p|g(x)|^p\Big)d\mu(x) \\ &= 2^p\|f\|_{L^p}^p + 2^p|\lambda|^p\|g\|_{L^p}^p\end{aligned}$$

dh. $\mathcal{L}^p(X)$ und $L^p(X)$ sind Vektorräume.

- Nach Lemma 14.3.11 gilt $\|f\|_{L^p}^p = \int\limits_X |f|^p d\mu = 0$ genau dann wenn $|f|^p = 0$ $\mu$-fast überall ist, was wegen Monotonie von $\begin{array}{ccc}[0,\infty[ & \to & [0,\infty[\\ x & \mapsto & x^p\end{array}$ äquivalent zu $|f| = 0$ $\mu$-fast überall ist, dh. zu $[f]_\sim = 0$ in $L^p(X)$. [1]

- Für $\lambda \in \mathbb{K}$ ist offenbar $\|\lambda f\|_{L^p} = \left(\int\limits_X |\lambda f(x)|^p d\mu(x)\right)^{\frac{1}{p}} = |\lambda|\|f\|_{L^p}$.

[1] Auf $\mathcal{L}^p(X)$ definiert $\|\cdot\|_{\mathcal{L}^p}$ nur eine Halbnorm, dh.

- $\|f\|_{\mathcal{L}^p} \ge 0$ für alle $f \in \mathcal{L}^p(X)$.
- $\|\alpha f\|_{\mathcal{L}^p} = |\alpha|\,\|f\|_{\mathcal{L}^p}$ für alle $f \in \mathcal{L}^p(X)$, $\alpha \in \mathbb{K}$ .
- $\|f+g\|_{\mathcal{L}^p} \le \|f\|_{\mathcal{L}^p} + \|g\|_{\mathcal{L}^p}$ für alle $f, g \in \mathcal{L}^p(X)$.

aber gegenüber einer Norm fehlt die Äquivalenz von $\|f\|_{L^p} = 0$ mit $f = 0$ in $L^p(X)$.

- $\|\cdot\|_{L^p}$ erfüllt wegen Monotonie des Integrals und der Dreiecksungleichung in $(\mathbb{K}, |\cdot|)$ die Dreiecksungleichung in $L^1(X)$. Für $p \in ]1, \infty[$ folgt mit $q = \frac{p}{p-1}$ aus der Hölder-Ungleichung:

$$\begin{aligned}
\|f+g\|_{L^p}^p &= \int\limits_X |f(x)+g(x)|^p d\mu(x) \\
&\leq \int\limits_X |f(x)||f(x)+g(x)|^{p-1} d\mu(x) + \int\limits_X |g(x)||f(x)+g(x)|^{p-1} d\mu(x) \\
&\leq (\|f\|_{L^p} + \|g\|_{L^p}) \left( \int\limits_X |f(x)+g(x)|^{(p-1)q} d\mu(x) \right)^{\frac{1}{q}} \\
&= (\|f\|_{L^p} + \|g\|_{L^p}) \|f+g\|_{L^p}^{p-1}
\end{aligned}$$

dh $\|f+g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}$. □

**Lemma 16.3.8.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $(f_n)_{n\in\mathbb{N}}$ eine Folge von nichtnegativen meßbaren Funktionen und $p \in [1, \infty[$, dann gilt*

$$\left\| \sum_{k=1}^{\infty} f_k \right\|_{L^p} = \lim_{n\to\infty} \left\| \sum_{k=1}^{n} f_k \right\|_{L^p} \leq \sum_{k=1}^{\infty} \|f_k\|_{L^p} \tag{16.3.10}$$

*Beweis.* Nach Korollar 14.1.17 ist auch $\sum\limits_{n=1}^{\infty} f_n$ eine nichtnegative meßbare Funktion, daher sind alle vorkommenden $\|\cdot\|_{L^p}$-Normen zumindest in $[0, \infty]$ definiert. Zum Beweis von (16.3.10) dürfen wir also oE. $\|f_k\|_{L^p} < \infty$ für alle $k \in \mathbb{N}$ voraussetzen. Für $f_1 + ... + f_n$ gilt nach der Minkowski-Ungleichung und wegen $\|f_k\|_{L^p} \geq 0$:

$$\|f_1 + ... + f_n\|_{L^p} \leq \sum_{k=1}^{n} \|f_k\|_{L^p} \leq \sum_{k=1}^{\infty} \|f_k\|_{L^p}. \tag{16.3.11}$$

Wegen $f_k \geq 0$ ist die Funktionenfolge $(f_1 + ... + f_n)_{n\in\mathbb{N}}$ (punktweise) monoton steigend und $0 \leq f_1(x) + ... + f_n(x) \leq f_1(x) + ... + f_n(x) + f_{n+1}(x) \nearrow \sum\limits_{k=1}^{\infty} f_k(x)$. Aus der Monotonie von $x \mapsto x^p$ folgt

$$0 \leq (f_1(x) + ... + f_n(x))^p \leq (f_1(x) + ... + f_n(x) + f_{n+1}(x))^p \nearrow \left( \sum_{k=1}^{\infty} f_k(x) \right)^p.$$

Nach dem Satz von der monotonen Konvergenz ist

$$\left\| \sum_{k=1}^{\infty} f_k \right\|_{L^p}^p = \sup \left\{ \left\| \sum_{k=1}^{n} f_k \right\|_{L^p}^p : n \in \mathbb{N} \right\} = \lim_{n\to\infty} \left\| \sum_{k=1}^{n} f_k \right\|_{L^p}^p$$

also folgt aus der Stetigkeit der $p$-ten Potenz:

$$\left\| \sum_{k=1}^{\infty} f_k \right\|_{L^p} = \sup \left\{ \left\| \sum_{k=1}^{n} f_k \right\|_{L^p} : n \in \mathbb{N} \right\} = \lim_{n\to\infty} \left\| \sum_{k=1}^{n} f_k \right\|_{L^p}. \tag{16.3.12}$$

Aus (16.3.11) und (16.3.12) folgt durch Bilden $\sup\limits_{n\in\mathbb{N}}$:

$$\left\| \sum_{k=1}^{\infty} f_k \right\|_{L^p} = \sup \left\{ \left\| \sum_{k=1}^{n} f_k \right\|_{L^p} : n \in \mathbb{N} \right\} \leq \sum_{k=1}^{\infty} \|f_k\|_{L^p}.$$

□

**Satz 16.3.9** (majorisierte Konvergenz in $\mathcal{L}^p(X)$)**.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum, $p \in [1, \infty[$ und $(f_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{L}^p(X)$, die $\mu$-fast überall punktweise konvergiert. Es sei $g : X \to [0, \infty] \in L^p(X)$ mit $|f_n(x)| \leq g(x)$ für alle $x \in X$, $n \in \mathbb{N}$. Dann definiert der punktweise Grenzwert $f(x) := \lim\limits_{n\to\infty} f_n(x)$ ein Element $f \in \mathcal{L}^p(X)$ mit $\|f_n - f\|_{L^p} \overset{n\to\infty}{\longrightarrow} 0$.*

*Beweis.* Nach Voraussetzung gibt es eine $\mu$-Nullmenge, oE. $N_1 \in \mathcal{A}$, so daß $\lim\limits_{n\to\infty} f_n(x)$ ($\in \widehat{\mathbb{R}}$ bzw. $\in \widehat{\mathbb{C}}$) für alle $x \in X \setminus N_1$ existiert. Da $g^p$ $\mu$-integrierbar ist, gibt es nach Lemma 14.3.11 eine $\mu$-Nullmenge, oE. $N_2 \in \mathcal{A}$ mit $g(x) < \infty$ für alle $x \in X \setminus N_2$. Damit ist durch

$$f(x) := \begin{cases} \lim\limits_{n\to\infty} f_n(x) & \text{für } x \in X \setminus (N_1 \cup N_2) \\ 0 & \text{für } x \in N_1 \cup N_2 \end{cases}$$

eine meßbare Funktion $f : X \to \mathbb{R}$ (bzw. $f : X \to \mathbb{C}$) definiert, so daß $f_n \overset{n\to\infty}{\longrightarrow} f$ $\mu$-fast überall und $|f| \leq g$ $\mu$-fast überall gelten. Wegen Monotonie des Integrals (von nichtnegativen meßbaren Funktionen) ist $\int\limits_X |f|^p d\mu \leq \int\limits_X g^p d\mu$, also $f \in \mathcal{L}^p(X)$. Setze $g_n := |f_n - f|^p$, so folgt

$$0 \leq g_n(x) \leq (|f_n(x)| + |f(x)|)^p \leq (|f(x)| + |g(x)|)^p$$

für alle $x \in X \setminus (N_1 \cup N_2)$, dh. die Funktionenfolge $(g_n)_{n\in\mathbb{N}}$ ist durch $h := (|f| + g)^p$ majorisiert. Das Lemma von Fatou ergibt:

$$\int\limits_X \left( \liminf_{n\to\infty} (h - g_n) \right) d\mu \leq \liminf_{n\to\infty} \int\limits_X (h - g_n) d\mu = \int\limits_X h d\mu - \limsup_{n\to\infty} \int\limits_X g_n d\mu. \tag{16.3.13}$$

Da $f_n \overset{n\to\infty}{\longrightarrow} f$ $\mu$-fast überall, folgt $g_n \overset{n\to\infty}{\longrightarrow} 0$ $\mu$-fast überall oder $h - g_n \overset{n\to\infty}{\longrightarrow} h$ $\mu$-fast überall, dh. $\liminf\limits_{n\to\infty} (h - g_n) = h$ $\mu$-fast überall. Die Majorante $h \in \mathcal{L}^1(X)$ der Folge $(h - g_n)_{n\in\mathbb{N}}$ ergibt nun $\int\limits_X h d\mu = \int\limits_X \left( \liminf\limits_{n\to\infty} (h - g_n) \right) d\mu$ nach dem Satz über majorisierte Konvergenz und damit folgt durch Einsetzen in (16.3.13): $\limsup\limits_{n\to\infty} \int\limits_X g_n d\mu \leq 0$; da aber $g_n \geq 0$ ist, folgt damit $\lim\limits_{n\to\infty} \int\limits_X g_n d\mu = 0$.

□

**Satz 16.3.10.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum, dann ist $(L^p(X), \|\cdot\|_{L^p})$ für $p \in [1, \infty]$ ein Banachraum. Ist $p \in [1, \infty[$ und $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $(L^p(X), \|\cdot\|_{L^p})$ mit Grenzwert $f \in L^p(X)$, dann gibt es eine Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$ von $(f_n)_{n\in\mathbb{N}}$, die $\mu$-fast überall gegen $f$ konvergiert.*

*Beweis.* Betrachte zunächst $p \in [1, \infty[$. Da $(f_n)_{n\in\mathbb{N}}$ in $(L^p(X), \|\cdot\|_{L^p})$ eine Cauchyfolge ist, können wir eine Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$ wählen mit

$$\|f_{n_{k+1}} - f_{n_k}\|_{L^p} < 2^{-k}$$

für jedes $k \in \mathbb{N}$ und fixieren einen Repräsentanten $f_{n_k} \in \mathcal{L}^p(X)$. Setze $g_k := f_{n_{k+1}} - f_{n_k}$ und $g := \sum_{k=1}^{\infty} |g_k|$, so sind für $(|g_k|)_{k\in\mathbb{N}}$ und $g$ die Voraussetzungen von Lemma 16.3.8 erfüllt, also ist

$$\|g\|_{L^p} \leq \sum_{k=1}^{\infty} \|g_k\|_{L^p} \leq \sum_{k=1}^{\infty} 2^{-k} = 1$$

und damit gilt für die $\mathcal{A}$-meßbare Funktion $g \geq 0$ auch $g \in \mathcal{L}^p(X)$. Nach Lemma 14.3.11 ist $g$ $\mu$-fast überall $[0, \infty[$-wertig, dh. die Reihe $\sum_{k=1}^{\infty} g_k(x)$ ist für $\mu$-fast alle $x \in X$ absolut konvergent. Da aber $\sum_{k=1}^{N} g_k(x) = f_{n_{N+1}}(x) - f_{n_1}(x)$ ist, konvergiert die Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$ $\mu$-fast überall und

$$|f_{n_{k+1}}(x)| = |g_1(x) + ... + g_k(x) + f_{n_1}(x)| \leq g(x) + |f_{n_1}(x)|.$$

Da die Potenzfunktion $x \mapsto x^p$ monoton ist, folgt

$$|f_{n_{k+1}}(x)|^p \leq (g(x) + |f_{n_1}(x)|)^p \leq 2^p g(x)^p + 2^p |f_{n_1}(x)|^p$$

und die rechte Seite gibt eine integrierbare Majorante für die Folge $(|f_{n_k}|^p)_{k\in\mathbb{N}}$ an. Nach dem Satz 16.3.9 über majorisierte Konvergenz in $\mathcal{L}^p(X)$ gilt für den $\mu$-fast sicheren Grenzwert $f(x) := \lim_{k\to\infty} f_{n_k}(x)$:

$$f \in \mathcal{L}^p(X) \quad \text{und} \quad \|f - f_{n_k}\|_p \overset{k\to\infty}{\longrightarrow} 0.$$

Da $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $L^p(X)$ ist, gibt es zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ mit $\|f_m - f_n\|_p < \frac{\varepsilon}{2}$ für alle $m, n \geq N(\varepsilon)$. Da $(f_{n_k})_{k\in\mathbb{N}}$ in $L^p(X)$ gegen $f$ konvergiert, gibt es ein $M = M(\varepsilon) \in \mathbb{N}$ mit $\|f_{n_k} - f\|_p < \frac{\varepsilon}{2}$ für alle $n_k \geq M(\varepsilon)$. Dann folgt aus der Dreiecksungleichung

$$\|f_n - f\|_p \leq \|f_n - f_{n_k}\|_p + \|f_{n_k} - f\|_p < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für alle $n \geq N(\varepsilon)$ und $n_k \geq \max\{M(\varepsilon), N(\varepsilon)\}$.

Wir kommen nun zum Fall $p = \infty$: Es sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $L^\infty(X)$, und wir wählen je einen Repräsentanten $f_n \in \mathcal{L}^\infty(X)$ von $f_n$. Dann sind nach Definition von

$\|\cdot\|_{\mathcal{L}^\infty}$ die Mengen $\{x \in X : |f_n(x)| > \|f_n\|_{\mathcal{L}^\infty}\}$, $n \in \mathbb{N}$ und $\{x \in X : |f_n(x) - f_m(x)| > \|f_n - f_m\|_{\mathcal{L}^\infty}\}$, $m, n \in \mathbb{N}$ jeweils $\mu$-Nullmengen und daher ist auch

$$N := \bigcup_{n\in\mathbb{N}} \{x \in X : |f_n(x)| > \|f_n\|_{\mathcal{L}^\infty}\} \cup \bigcup_{m,n\in\mathbb{N}} \{x \in X : |f_n(x) - f_m(x)| > \|f_n - f_m\|_{\mathcal{L}^\infty}\}$$

eine $\mu$-Nullmenge. Da $|f_m(x) - f_n(x)| \leq \|f_m - f_n\|_{\mathcal{L}^\infty}$ für jedes $x \in X\backslash N$ erfüllt ist, konvergiert $(f_n \mathbf{1}_{X\backslash N})_{n\in\mathbb{N}}$ gleichmäßig gegen den punktweisen Grenzwert

$$\begin{array}{rcl} f\mathbf{1}_{X\backslash N} : X\backslash N & \to & \mathbb{C} \\ x & \mapsto & \lim\limits_{n\to\infty} f_n(x) \end{array}$$

und es ist $f\mathbf{1}_{X\backslash N} \in \mathcal{L}^\infty(X)$ und $\lim\limits_{n\to\infty} \|f_n - f\mathbf{1}_{X\backslash N}\|_{\mathcal{L}^\infty}) = 0$ □

**Satz 16.3.11.** *Es sei $(X, \mathcal{A}, \mu)$ ein Maßraum, dann wird $L^2(X)$ durch das Skalarprodukt*

$$\begin{array}{rcl} \langle\cdot,\cdot\rangle_{L^2} : L^2(X) \times L^2(X) & \to & \mathbb{C} \\ (f, g) & \mapsto & \langle f, g\rangle_{L^2} := \int\limits_X \overline{f(x)} g(x) d\mu(x) \end{array}$$

*zu einem Hilbertraum.*

*Beweis.* Sind $f, g \in L^2(X)$, dann ist wegen $1 = \frac{1}{2} + \frac{1}{2}$ die Hölder Ungleichung anwendbar und ergibt:

$$\left|\int\limits_X \overline{f(x)} g(x) d\mu(x)\right| \leq \int\limits_X |f(x)| \cdot |g(x)| d\mu(x) = \| \, |f| \cdot |g| \, \|_1 \leq \|f\|_2 \|g\|_2$$

dh. $\int\limits_X \overline{f} g d\mu \in \mathbb{C}$ ist wohldefiniert und damit ist $\langle\cdot,\cdot\rangle_{L^2}$ nach den Rechenregeln für das Integral eine Sesquilinearform auf $L^2(X)$. Weil $\|\cdot\|_{L^2}$ eine Norm auf $L^2(X)$ definiert ist wegen $\|f\|_{L^2}^2 = \langle f, f\rangle_{L^2}$ die Sesquilinearform $\langle\cdot,\cdot\rangle_{L^2}$ positiv definit. $L^2(X)$ versehen mit $\|\cdot\|_{L^2}$ ist nach Satz 16.3.10 vollständig. □